

Top-down control of visual attention and its reflection in human visual cortex

Citation for published version (APA):

Peters, J. C. (2007). *Top-down control of visual attention and its reflection in human visual cortex*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht.
<https://doi.org/10.26481/dis.20071024jp>

Document status and date:

Published: 01/01/2007

DOI:

[10.26481/dis.20071024jp](https://doi.org/10.26481/dis.20071024jp)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Top-down control of visual attention and its reflection in human visual cortex

Judith Carolien Peters

Colophon

Cover design by Mirko Reithler & Judith Peters
Production: Datawyse | Universitaire Pers Maastricht
Copyright © Judith C. Peters, Maastricht 2007
All rights reserved

ISBN 978-90-5278-672-8

Top-down control of visual attention and its reflection in human visual cortex

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,
op gezag van de Rector Magnificus, prof. mr. G.P.M.F. Mols,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op woensdag 24 oktober 2007 om 16.00 uur

door

Judith Carolien Peters

Geboren op 9 september 1979 te Leiden



Promotor:

Prof. dr. R. Goebel

Copromotor:

Prof. dr. P.R. Roelfsema (Nederlands Instituut voor Neurowetenschappen
[KNAW]; Vrije Universiteit)

Beoordelingscommissie:

Prof. dr. C. Kemner (voorzitter)

Dr. A.T. Sack

Prof. dr. F. di Salle

Prof. dr. J. Theeuwes (Vrije Universiteit)

Het in dit proefschrift beschreven onderzoek werd mede mogelijk gemaakt door een subsidie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO, nr. 402-01-632) toegekend aan prof. dr. R. Goebel.

Voor mijn moeder

Table of contents

Chapter 1	General introduction	p.9
Chapter 2	Remembered but unused: the accessory items in working memory that do not guide attention	p.27
Chapter 3	Neural correlates of attentional guidance by items in visual working memory	p.47
Chapter 4	Scaling the zoom-lens of attention	p.69
Chapter 5	The neural mechanisms of scaling the attentional zoom-lens	p.101
Chapter 6	Visual attention to multiple stimuli: To shift or to divide?	p.123
Chapter 7	Reflections of complex spatial attention distributions in human visual cortex	p.167
Chapter 8	Summary	p.191
	Samenvatting	p.197
	Acknowledgments	p.203
	Curriculum Vitae	p.205

Chapter 1

General introduction

From the moment we open our eyes, our brain is continuously bombarded with visual information, which would cause an information overload if everything would be processed simultaneously. Visual attention enables us to select information that is most relevant for further processing and to ignore irrelevant, distracting information. Considering the complexity of the visual environment in today's society, it is of utmost importance that relevant information is efficiently and reliably selected from the avalanche of incoming percepts. For example, while driving, other cars, pedestrians, traffic lights and stop signs need to be attended, while neon advertising signs need to be ignored. Likewise, watching your children on a crowded play ground or operating air traffic control equipment, asks for an optimal allocation of one's limited attentional resources. However, when you picture yourself in such a demanding situation (e.g., driving downtown towards an intersection packed with pedestrians, bicyclers, cars, and traffic signs), you may wonder how effectively your attention can be deployed. Can you simultaneously attend multiple objects? Can you ignore all information that you are currently not interested in? The studies presented in this thesis aimed to explore how proficiently attention can be voluntarily allocated in a range of situations. Moreover, the neural mechanisms enabling this 'top-down controlled' selection as well as the effects of attention on sensory representations in visual cortex were studied using neuroimaging methods such as functional magnetic resonance imaging (fMRI) and electro-encephalography (EEG). To introduce the topic in more detail, a short (and selective) background on the (neural) mechanisms of visual attention and the used neuro-imaging methods will be provided below, followed by a more profound outline of the studies discussed in the subsequent chapters.

The mechanisms of covert visual attention

It is evident that attention has a beneficial effect on performance, which has been also empirically demonstrated in numerous studies. Stimuli at attended positions are processed (Carrasco, & McElree, 2001) and detected (e.g., Posner, Snyder, & Davidson, 1980) faster, are discriminated better (e.g., Downing, 1988) and are recognized more accurately (e.g., Van der Heijden, Schreuder, & Wolters, 1985) than stimuli at unattended positions. Moreover, it has been shown that attention can strengthen the representation of an attended stimulus by increasing contrast sensitivity (e.g., Carrasco, Penpeci-Talgar, & Eckstein, 2000; Reynolds, Pasternak, & Desimone, 2000) or spatial resolution (Yeshurun & Carrasco, 1998, 1999, 2000), although remarkably, this can under some circumstances result in decreased performance (Ling & Carrasco, 2006; Yeshurun & Carrasco, 1998). Interestingly, recent findings suggest that the signal enhancement of an attended stimulus is accompanied by a change in the phenomenological experience of that stimulus. For example, attending to a stimulus increases the apparent stimulus contrast (Carrasco, Ling, & Read, 2004) and spatial frequency (Gobell & Carrasco, 2005).

Moreover, the subjective experience of flicker rate (Montagna & Carrasco, 2006), color saturation (Fuller, & Carrasco, 2006), and motion coherence (Liu, Fuller, & Carrasco, in press) can be altered by attention. However, in addition to signal enhancement of attended stimuli, attention can also reduce the influence of unattended stimuli (e.g., Doshier, & Lu, 2000a, 2000b). Although the consequences of directing attention to a stimulus appear to be straightforward, the mechanisms generating such effects are not as clearly defined. Covert spatial attention is often compared to a ‘spotlight’ (Posner, 1980), a ‘beam’ of attention (with a fixed size and shape) that can move across the visual field, independent of eye movements, thereby ‘illuminating’ selected locations in which processing is facilitated by attention. However, experimental results indicate that the actual operation of spatial attention has some characteristics that would not be predicted by the spotlight metaphor (for review, see Cave & Bichot, 1999). For example, the size of the beam is not fixed, but can be adjusted in accordance to specific task requirements (e.g., LaBerge, 1983). This finding has led to the introduction of the ‘zoom-lens’ metaphor (Eriksen & St. James, 1986), which is closely related to the spotlight metaphor. The zoom-lens model assumes that the attentional focus can be constricted or expanded to match the spatial scale of a relevant region, accompanied by a tradeoff between the size of the attentional focus and processing efficiency. The flexibility of the zoom-lens’ size and the negative correlation between size and processing efficiency have been confirmed by a range of studies (Castiello & Umiltà, 1990; Luo, Greenwood, & Parasuraman, 2001; McMains & Somers, 2005; Müller, Bartelt, Donner, Villringer, & Brandt, 2003). Other findings have revealed a gradual decrease in attention effects from the center towards the edges of the attended area, suggesting a ‘gradient’ of attentional enhancement (Downing & Pinker, 1985; LaBerge & Brown, 1989; Mangun & Hillyard, 1987; Shulman, Sheehy, & Wilson 1986). Moreover, not only the size of the beam, but also the shape of the beam appears to be flexible. That is, although the beam can have a circular shape (Eriksen & St. James, 1986) as suggested by the spotlight metaphor, attention can also be deployed in an annular configuration (see Chapter 7 of this thesis; Egly & Homa, 1984; Juola, Bouwhuis, Cooper, & Warner, 1991; Müller & Hübner, 2002). Furthermore, although some initial studies (Shulman, Remington, & MacLean, 1979; Tsal, 1983) suggested that the attentional beam is shifted from one location to another in an analog manner (i.e., covering other locations along the way) in line with the behavior of a spotlight, later studies (e.g., Chastain, 1992a, 1992b; Murphy & Eriksen, 1987; Sperling & Weichselgartner, 1995) mainly provided support for a discretely moving beam (i.e., ‘jumping’ from the first to the second location). In addition, several psychophysical (Scharlau, 2004; Awh & Pashler, 2000; Bichot et al., 1999; Hahn & Kramer, 1998; Kramer & Hahn, 1995; Castiello & Umiltà, 1992) and neuroimaging (Müller, Malinowski, Gruber, & Hillyard, 2003; McMains & Somers, 2004) studies (for review, see Chapter 6 of this thesis) found evidence that attention can be divided over non-contiguous regions, which is in

contradiction with the unitary focus assumed by the spotlight metaphor. Finally, the concept of a spotlight (or zoom-lens, or gradient) suggests that allocation of visual attention is solely mediated by spatial aspects. Conversely, several studies using a range of methods (e.g., Duncan, 1984; Vecera & Farah, 1994; Roelfsema, Lamme, & Spekreijse, 1998; O'Craven, Downing, & Kanwisher, 1999; for review, see Scholl, 2001) have shown that attention can also be directed to a perceptual group or to an object, even if objects spatially overlap. However, the extent to which object-based attention is independent of spatial representations is still under debate (Vecera and Farah, 1994; for reviews, see Cave & Bichot, 1999; Mozer & Vecera, 2005).

Attention can be directed to stimuli in an endogenous or exogenous manner (Posner, 1980). Endogenous (or goal-directed) attention is regulated by 'top-down' factors such as the goals and the expectations of the observer. In contrast, 'bottom-up' processes control exogenous (or stimulus-driven) attention, leading to 'attentional capture' by salient stimuli such as color singletons (e.g., Theeuwes, 1991, 1992, 1994) or abrupt visual onsets (e.g., Theeuwes, 1994; Yantis & Jonides, 1984). Thus, whereas attention can be voluntarily directed to parts of the visual scene, salient stimuli can capture attention, causing an involuntary shift of attention to these stimuli. The extent to which top-down control can override attentional capture by salient distractors during the initial stages of visual processing, has been intensively debated over the last decades (for reviews, see Rauschenberger, 2003; Ruz, & Lupiáñez, 2002). In this debate, accounts range from complete top-down control (e.g., Folk, Remington, & Johnston, 1992) to a complete lack of top-down control (e.g., Theeuwes, 1991, 1992, 1994), according to whether salient distractors were able to capture attention in the evaluated visual search experiments or not. An explanation for these divergent findings has been proposed by Theeuwes (2004), who pointed out that in studies that did not reveal attentional capture by salient distractors, (small clusters of) items in the search array were most likely examined consecutively with a small attentional window (e.g., Jonides & Yantis, 1988), whereas a broad attentional window was applied to the search array in studies that did observe attentional capture by salient distractors. If salient items can only capture attention when they are located within the attentional window (since the underlying salience computations are largely restricted to this window), salient distractors would less often capture attention when the attentional window is relatively small. Recent psychophysical findings (Belopolsky, Zwaan, Theeuwes, & Kramer, in press) support this hypothesis, suggesting that although the size of the attentional window is under top-down control, no top-down control can be exerted over bottom-up processes within the attentional window. However, although bottom-up factors should be taken into account, this thesis focuses on the mechanisms of top-down controlled voluntary attention rather than bottom-up saliency-based attention.

Neuroimaging methods

Although the fMRI method (for reviews, see e.g. Amaro & Barker, 2006; Bandettini, Birn, & Donahue, 2000) has only recently been developed, it already contributed greatly to our understanding of human brain function. Using fMRI, neural activity can be measured simultaneously in the whole brain with a high spatial resolution, in a non-invasive manner. However, since fMRI measures the vascular response caused by neural activation instead of the neuronal processing itself, it provides only an indirect measure of neural activity. The majority of fMRI studies (including the studies presented in this thesis) use the blood-oxygen level-dependent (BOLD) contrast mechanism (e.g., Ogawa et al., 1990a, 1990b) indexing the ratio between oxy- and deoxyhemoglobin, which is higher in activated areas since the oxygen supplied by the increased regional blood flow exceeds the oxygen demand. Consequently, the fMRI signal is a delayed and temporally as well as spatially smoothed response compared to the underlying neural activity, leading to a relatively low temporal resolution of the method. Another neuroimaging method that is also non-invasive, but which can measure neural activity with high temporal resolution, is the Event-Related Potential (ERP) method (for review, see e.g. Kutas & Dale, 1997). By means of EEG, voltage fluctuations that are generated by the synchronous activation of neural populations can be measured on the scalp. From EEG data, ERPs can be extracted by averaging phasic changes in the ongoing EEG, which are time-locked to a repeated event (such as the presentation of a stimulus). Since the voltage fluctuations are instantaneously coupled to changes in neural activity, the ERP method enables measuring the time-course of ongoing cognitive processes with high temporal resolution. However, since the spatial resolution is rather low, fMRI is more suitable for the localization of neural sources.

Most fMRI studies have used ‘block designs’, in which responses to blocks of successively presented stimuli are studied. Since the responses to the individual stimuli are integrated over time in the obtained responses, the time course of neural activity evoked by a single event cannot be evaluated in a block design. However, the subsequently developed ‘event-related designs’ (for review, see Rosen, Buckner, & Dale, 1998), in which trials are presented separately rather than in clusters, do permit the study of transient hemodynamic variations related to individual events (by time-locked averaging of responses similar to the ERP method). Since these designs allow applications that are often used in cognitive psychology, such as the randomization of conditions or post-hoc categorization of trials based on behavioral performance, the introduction of event-related procedures greatly increased the possibilities to study cognitive processes using fMRI. Because the hemodynamic or BOLD response only returns to baseline after about 20 seconds, the first event-related studies presented events with a rather slow rate, to minimize the overlap between BOLD responses to succeeding events (e.g., Buckner et al., 1996). However, additional studies showed that BOLD

responses to stimuli presented in rapid succession (with inter-stimulus intervals as short as 2 s) sum in an approximately linear fashion (Boynton, Engel, Glover, & Heeger, 1996; Dale & Buckner, 1997), indicating that BOLD responses can be modeled as a linear system. Therefore, responses to individual items can still be estimated even if the stimulus presentation rate is considerably increased, given that the inter-stimulus interval is sufficiently jittered between trials (Dale, 1999; Miezin, Maccotta, Ollinger, Petersen, & Buckner, 2000). Moreover, by using deconvolution techniques (Serences, 2004; Glover, 1999; see Chapter 3 & 5 of this thesis for applications), which independently estimate each data point of the hemodynamic response function, the time-course of the BOLD responses can be evaluated with minimal contamination of neural responses to temporally adjacent events. The possibility to disentangle overlapping responses to rapidly succeeding events (see Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000; Woldorff et al., 2000 for additional methods), enabled the study of isolated mental operations in tasks without the need for too long intervals between events that could change the nature of the task. For example in cueing paradigms (such as the Posner task; see next section), cue-induced orienting responses can be investigated in isolation from subsequent stimulus processing responses, without the need for a (compared to psychophysical standards) extraordinarily long cue-delay.

Functional neuroanatomy of covert visual attention

The fronto-parietal network involved in controlling attentional orienting

Based on lesion studies in patients with visuospatial neglect, first attempts to localize the neural basis of attentional orienting were made (e.g., Mesulam, 1981). However, the subsequent advent of fMRI provided the means to non-invasively study the human functional neuroanatomy of attention in more detail in non-clinical populations. The majority of fMRI studies have focused on the mechanisms controlling endogenous orienting of visual spatial attention. Most of these studies used a (modified) version of the cueing paradigm developed by Posner (Posner, 1980; Posner & Cohen, 1984), in which a centrally presented symbolic cue (such as an arrow) instructs participants to direct attention to a location in the periphery, where a subsequent target will most likely be presented. Cue-induced endogenous orienting of attention enhances activation in a dorsal fronto-parietal network (Corbetta & Shulman, 2002), including the intraparietal sulcus (IPS) and the superior parietal lobule (SPL) in the posterior parietal cortex (PPC), and the frontal eye field (FEF) located in the superior posterior frontal cortex (Figure 1), as core regions (Corbetta, et al., 2000; Hopfinger, Buonocore, & Mangun, 2000). The different areas constituting the dorsal fronto-parietal network appear to mediate distinct attentional control processes. The FEF and IPS show a sustained increased response throughout a delay period following the cue (Corbetta, Kincade, & Shulman, 2002), indicating that these areas are involved in

maintaining a peripheral attention focus. In contrast, a study by Yantis and colleagues (2002) revealed that activity in the right SPL (and inferior parietal lobule; IPL) is only transiently enhanced upon cue presentation. Furthermore, the results showed that, unlike the IPS that contains topographic maps of attention (Silver, Ress, & Heeger, 2005), the SPL did not differentiate between left- and rightward shifts of attention. These findings suggested that the SPL plays a role in attention shifting, although without specifying to which location attention should be shifted. Note that Yantis and colleagues (2002) did not use Posner's cueing paradigm, but instead used a task in which participants monitored one of two letter streams (presented left and right of fixation, each surrounded by three distractor letter streams) for a digit, which indicated whether attention should be maintained ('hold') on the current stream or whether attention should be shifted ('shift') to the other stream. By contrasting shift and hold conditions, areas involved in the actual shifting (and disengaging) of attention between two locations could be studied in isolation of other attention processes. Subsequently, Yantis and colleagues applied this paradigm to study attention shifts between objects (Serences, Schwarzbach, Golay, Courtney, & Yantis, 2004), between features (Liu, Slotnick, Serences, & Yantis, 2003), and between modalities (Shomstein & Yantis, 2004). Interestingly, similar shift-related transient activations within the PPC were revealed across these studies, although the specific spatial activation patterns depended on the type of attention shift. The recruitment of the (medial superior) PPC in spatial as well as non-spatial attention shifts, in combination with the lack of sensitivity to the direction of spatial attention shifts (Yantis et al., 2002), have led to the suggestion that this transient shift-related activation is rather a more abstract 'reset' signal than a signal conveying information concerning the target of an attention shift (Serences & Yantis, 2006). On the other hand, results of a recent study (Shomstein & Behrmann, 2006) suggested that the strength of transient activity in the left medial superior PPC is modulated by the context in which spatial attention shifts occur (i.e., shifting between positions located in the same object or in different objects), motivating further research to determine the exact role of transient PPC activity in attention shifting.

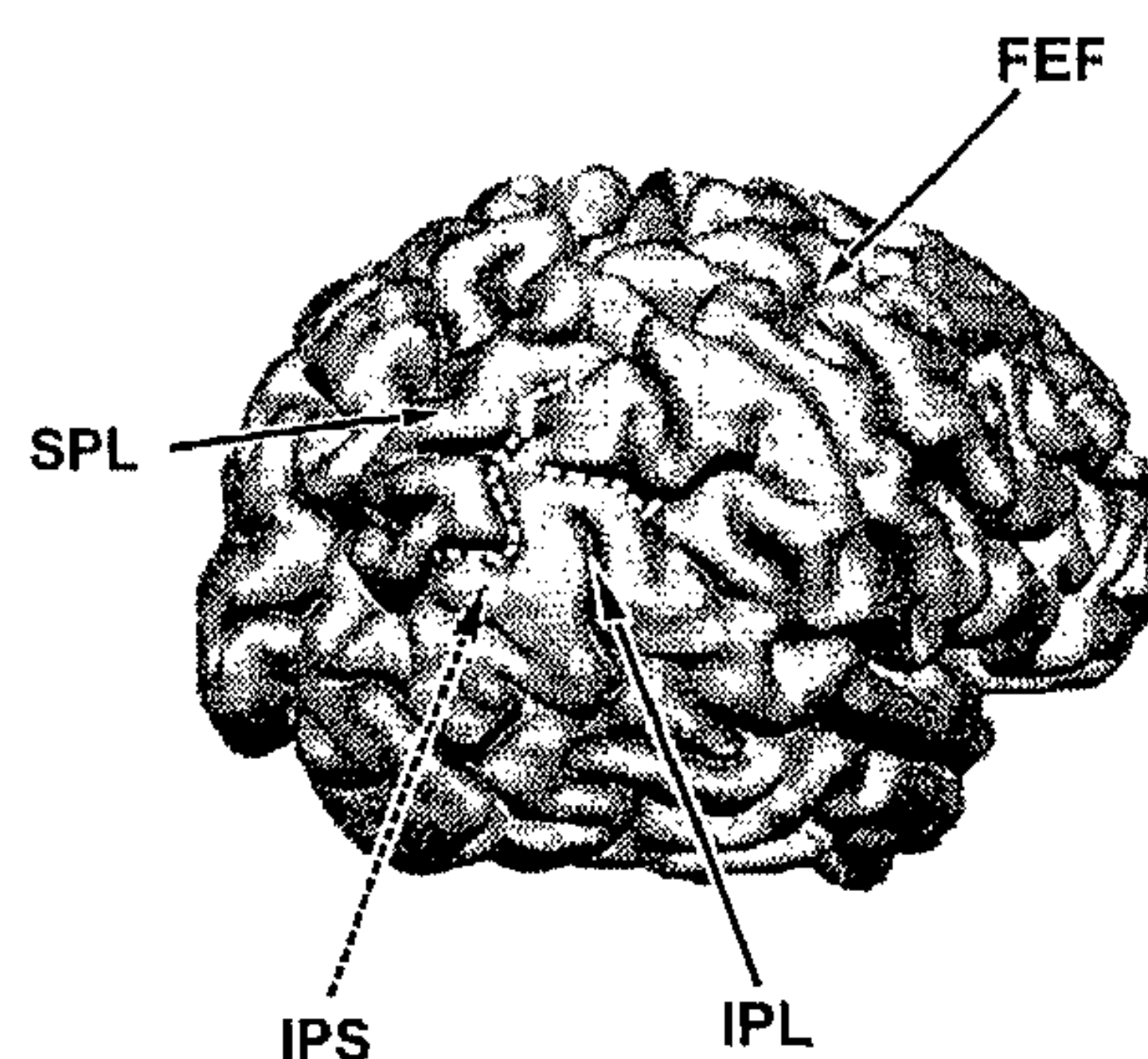


Figure 1. Frontal and parietal areas involved in voluntary attentional control shown on a reconstructed cortical surface of an individual brain.

In addition to endogenous orienting of attention, exogenous or stimulus-driven orienting can also be investigated using the Posner cueing paradigm. Instead of endogenously cueing attention to the periphery by a central symbol, briefly presenting a cue in the periphery causes an exogenous attention shift towards this cue. To investigate whether exogenous and endogenous orienting of attention were mediated by the same mechanisms, several fMRI studies compared neural activity in these two versions of the Posner paradigm revealing a large overlap between the exogenous and endogenous orienting network (Kim et al., 1999; Rosen et al., 1999; Peelen, Heslenfeld, & Theeuwes, 2004). However, activity related to cue-induced orienting of attention was not isolated from subsequent target processing in these studies, which could have obscured dissociations between exogenous and endogenous orienting networks. For example, a study by Corbetta and colleagues (2000), in which cue- and target-related processes were investigated separately, showed that IPS was more strongly activated during cue-induced endogenous orienting, whereas the precuneus and the right temporoparietal junction (TPJ) responded more during subsequent target processing, especially when the target occurred at an unexpected location. This latter result agrees with the suggestion that TPJ is part of a right-lateralized ventral fronto-parietal network that detects behaviorally relevant stimuli, particularly when they are unexpected or salient (Corbetta & Shulman, 2002). A recent study by Kincade and colleagues (2005) underscored the importance of behavioral relevance for TPJ involvement in exogenous orienting, by showing that TPJ activity was increased by orienting to targets at unattended locations, but not to irrelevant color singletons. Interestingly, task-irrelevant distractors are able to modulate TPJ responses when their color matches the target color, indicating that TPJ activity is sensitive to 'contingent' (Folk et al., 1992) attentional capture (Serences et al., 2005). The study of Kincade and colleagues (2005) further showed that exogenous orienting to singletons activated, next to visual areas specialized in color processing, a part of the dorsal fronto-parietal network (the left FEF in particular) involved in endogenous orienting. By investigating cue-induced orienting in isolation of subsequent target processing, this study substantiated earlier findings (Kim et al., 1999; Rosen et al., 1999; Peelen et al., 2004) indicating that exogenous and endogenous orienting are mediated by highly similar networks. Moreover, recent findings indicate that IPS and FEF responses are modulated by both endogenous as well as exogenous orienting to the contralateral (compared to the ipsilateral) visual field (Serences & Yantis, 2007). The sensitivity to both stimulus-driven as well as goal-directed factors of attention in combination with the spatial selectivity of these areas, suggest that the IPS and FEF contain spatial attentional 'priority maps' (Serences & Yantis, 2006, 2007; related to the concept of a salience map; e.g., Fecteau & Munoz, 2006), which would be in line with results of single-cell recordings in monkey lateral intraparietal area (LIP; Bisley & Goldberg, 2003; Gottlieb, Kusunoki, & Goldberg, 1998; Kusunoki, Gottlieb, & Goldberg, 2000) and FEF (e.g., Thompson & Bichot 2005).

Effects of attention in the visual cortex

Attentional orienting processes are reflected in the visual cortex. Neural responses in the SPL and FEF (and supplementary eye field) bias activity in the visual cortex (Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999), in which retinotopic regions representing the cued location show enhanced responses both in presence (Brefczynski & DeYoe, 1999; Somers, Dale, Seiffert, & Tootell, 1999; Tootell et al., 1998) as well as in absence (Kastner et al., 1999; Ress, Backus, & Heeger, 2000) of visual stimuli. The close match between the cortical topography of attentional enhancements and the retinotopic representations of the visual field, in combination with the relatively high spatial resolution of fMRI, allows a detailed study of the spatial attention distribution in various settings. By carefully examining the neural activation patterns in the visual cortex, support has been found for divided attention (McMains & Somers, 2004), the zoom-lens metaphor (McMains & Somers, 2005; Müller et al., 2003), attentional inhibition (Müller, Kleinschmidt, 2004; Slotnick, Schwarzbach, & Yantis, 2003), and interactions between object- and space-based attention have been observed (Müller & Kleinschmidt, 2003).

Outline of the thesis

The current thesis aims to investigate the neural structures involved in goal-directed selection of visual information, and the reflection of these top-down attentional control processes on sensory representations in the visual cortex. Moreover, to explore how proficiently goal-directed attention can be deployed in a range of situations, the studies presented in this thesis used various tasks that all required an efficient deployment of the limited attentional resources (e.g., Shaw & Shaw, 1978). These studies were part of three main projects, with each project focusing on a different aspect of voluntarily attentional control. The first project examined whether task-irrelevant stimuli that match information in working memory (WM) can be ignored. The second project studied how efficiently the attentional focus can be scaled. Finally, the third project explored the limits of dividing attention in a complex configuration. Or, to stick to the example of driving downtown towards a crowded intersection mentioned above, the projects would respectively answer questions such as: If you have someone on your mind that you are about to meet and you are searching for the street name signs (to check whether you are heading in the right direction), will you notice this person in the crowd at the corner of the intersection? And can you discretely switch between attending the entire intersection and attending the small street name signs? Moreover, would it be possible to divide your attention over the traffic and the signs? The specific aspects of each of the projects will be discussed in more detail below.

Chapter 2 & 3: Can items that are held online, be ignored when encountered in a visual scene?

WM appears to play an important role in the top-down guidance of attention, as after all, the goals that should direct our attention have to be kept in memory. According to the influential biased competition model (Desimone & Duncan, 1995), incoming visual signals compete with each other for further selective processing and this competition can be biased in favor of inputs matching WM content. Converging evidence from single-cell recordings has supported this model by showing that prefrontal structures involved in maintaining visual information are able to provide top-down feedback to visual areas (e.g., Fuster, Bauer, & Jervey, 1985; Tomita, Ohbayashi, Nakahara, Hasegawa, & Miyashita, 1999), resulting in a competitive advantage for matching visual inputs to be selected for attentional processing (Chelazzi, Duncan, Miller, & Desimone, 1998; Chelazzi, Miller, Duncan, & Desimone, 1993). Since multiple items can be held in WM (Luck & Vogel, 1997), an important question is whether all these items guide attention to matching stimuli or only the ones that are relevant for the task at hand. Recently, several behavioral studies (Olivers, Meijer, & Theeuwes, 2006; Woodman & Luck, in press; Houtkamp & Roelfsema, 2006; Soto et al, 2005; Downing & Dodds, 2004) investigated this question using a visual search task in which participants searched for a target object, while maintaining a second object in WM that would function as target in a subsequent task. Mixed results were reported concerning the capability of this second WM representation to guide attention, therefore warranting further research. Since behavioral performance only reflects the outcome of information processing, we used the high temporal resolution of the ERP method to examine the influence of WM content on the different stages of stimulus processing (e.g., Sanders, 1990) in a comparable visual search paradigm (*Chapter 2*). An extended version of this paradigm was employed in a subsequent fMRI study (*Chapter 3*). Here, visual searches across object categories were additionally included to study category-specific effects in specialized higher visual areas.

Chapter 4 & 5: Can the spatial attention focus be scaled discretely?

A vast amount of psychophysical (for review, see e.g. Cave & Bichot, 1999), and neuroimaging (see previous section) studies have examined the mechanisms of spatial attention shifts in humans when searching for (or preparing for the appearance of) a target. However during visual search, it is not only important to shift, but also to scale the attentional focus (Greenwood & Parasuraman, 1999; Hopf et al., 2006). Nevertheless, not much is known about the mechanisms underlying the scaling of the attentional focus. Therefore, a series of four psychophysical studies was performed (discussed in *Chapter 4*) to extend our knowledge concerning the key characteristics of the attentional scaling mechanism. For example, the influence of the ‘direction’ (i.e., zooming-in versus zooming-out) and the extent (i.e., the degree of change in size) of scaling on

attentional scaling processes were explored. The primary interest of these studies was to estimate whether the attentional focus can be discretely scaled from a large to a small size, in order to 'skip' distracting information that is presented at an intermediate level of scaling. In addition, a related fMRI study (*Chapter 5*) aims to elucidate the neural mechanisms involved in scaling the attentional focus and explores modulations of neural activity pointing to either discrete or continuous scaling of the attentional focus.

Chapter 6 & 7: Can attention be deployed in a complex configuration, in which multiple non-contiguous regions are attended?

As previously discussed, attentional configurations (i.e., the constellation of attended and non-attended areas) can be more complex than a single focused peak of attention as suggested by the spotlight or zoom-lens metaphor. For example, attention can be deployed in an annular fashion. Moreover, although intensively debated (a review of this debate is provided in *Chapter 6*), it appears to be possible to attend multiple non-contiguous regions in parallel. This flexibility of the attentional distribution is very beneficial, given that due to limited attentional resources, a tradeoff exists between the size of the attended region and processing efficiency (e.g., Eriksen & St. James, 1986). Accordingly, it would be highly efficient if valuable attentional resources are not wasted on irrelevant intervening areas, as would be the case for a unitary 'spotlight' or 'zoom-lens' encompassing an entire contiguous region. Such an efficient implementation would not only increase the amount of attention that can be deployed to relevant areas, but would also provide the possibility to ignore distracting information presented in irrelevant areas which otherwise could interfere with the execution of a task. However, not much is known about the degree of complexity in which spatial attention can be configured. Therefore, the studies discussed in *Chapter 7* used a paradigm in which attentional configurations that differed in complexity could be compared, while sensory stimulation was kept constant. Both the behavioral and neural counterparts of allocating attention in a complex configuration (e.g., two concentric but non-contiguous regions) were explored in, respectively, two psychophysical studies and a corresponding fMRI study.

Finally, the results of the studies presented in this thesis will be discussed and summarized in *Chapter 8*, which will also provide suggestions for new lines of research.

References

- Amaro, E. Jr., & Barker, G. J. (2006). Study design in fMRI: basic principles. *Brain and Cognition*, 60(3), 220-232.
- Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance*, 26(2), 834-846.
- Bandettini, P. A., Birn, R. M., & Donahue, K. M. (2000). Functional MRI: Background, methodology, limits, and implementation. In: J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (pp. 978-1014). University Press: Cambridge.
- Belopolsky, A. V., Zwaan, L., Theeuwes, J., & Kramer, A. F. (in press). The size of an attentional window modulates attentional capture by color singletons. *Psychonomic Bulletin & Review*.
- Bisley, J. W., & Goldberg, M. E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. *Science*, 299, 81-86.
- Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *The Journal of Neuroscience*, 16, 4207-4221.
- Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nature Neuroscience*, 2(4), 370-374.
- Buckner, R. L., Bandettini, P. A., O'Craven, K. M., Savoy, R. L., Petersen, S. E., Raichle, M. E., & Rosen, B. R. (1996). Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences USA*, 93(25), 14878-14883.
- Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, 7(3), 308-313.
- Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proceedings of the National Academy of Sciences USA*, 98, 5363-5367.
- Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity across the CSF: Support for signal enhancement. *Vision Research*, 40(10-12), 1203-1215.
- Castiello, U., & Umiltà, C. (1990). Size of the attentional focus and efficiency components of processing. *Acta Psychologica*, 73, 195-209.
- Castiello, U., & Umiltà, C. (1992). Splitting focal attention. *Journal of Experimental Psychology: Human Perception and Performance*, 18(3), 837-848.
- Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: beyond a spotlight model. *Psychonomic Bulletin & Review*, 6(2), 204-223.
- Chastain, G. (1992a). Analog versus discrete shifts of attention across the visual field. *Psychological Research*, 54, 175-181.
- Chastain, G. (1992b). Time-course of sensitivity changes as attention shifts to an unpredictable location. *Journal of General Psychology*, 119, 105-111.
- Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, 80, 2918-2940.
- Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, 363, 345-347.
- Corbetta, M., Kincade, J. M., Shulman, G. L. (2002). Neural systems for visual orienting and their relationships to spatial working memory. *Journal of Cognitive Neuroscience*, 14(3), 508-523.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3, 201-215.

- Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, 3, 292-297.
- Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, 8, 109-114.
- Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trial using fMRI. *Human Brain Mapping*, 5, 1-12.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193-222.
- Doshier, B. A., & Lu, Z.-L. (2000a). Mechanisms of perceptual attention in precuing of location. *Vision Research*, 40, 1269-1292.
- Doshier, B. A., & Lu, Z.-L. (2000b). Noise exclusion in spatial attention. *Psychological Science*, 11(2), 139-146.
- Downing, C. J. (1988). Expectancy and visual-spatial attention: effects on perceptual quality. *Journal of Experimental Psychology: Human Perception and Performance*, 14(2), 188-202.
- Downing, P. E., & Dodds, C. M. (2004). Competition in visual working memory for control of search. *Visual Cognition*, 11(6), 689-703.
- Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. Posner & O. Martin (Eds.), *Attention and Performance XI* (pp. 171-187). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 501-517.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96(3), 433-458.
- Egely, R., & Homa, D. (1984). Sensitization of the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 778-793.
- Eriksen, C. W., & James, D. S. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, 40(4), 225-240.
- Fecteau, J. H., & Munoz, D. P. (2006). Saliency, relevance, and firing: a priority map for target selection. *Trends in Cognitive Sciences*, 10(8), 382-390.
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance*, 18, 1030-1044.
- Fuller, S., & Carrasco, M. (2006). Exogenous attention and color perception: performance and appearance of saturation and hue. *Vision Research*, 46(23), 4032-4047.
- Fuster, J. M., Bauer, R. H., & Jervey, J. P. (1985). Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Research*, 330(2), 299-307.
- Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage*, 9(4), 416-429.
- Gobell, J., & Carrasco, M. (2005). Attention alters the appearance of spatial frequency and gap size. *Psychological Science*, 16(8), 644-651.
- Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, 391(6666), 481-484.
- Greenwood, P. M., & Parasuraman, R. (1999). Scale of attentional focus in visual search. *Perception & Psychophysics*, 61(5), 837-859.
- Hahn, S., & Kramer, A. (1998). Further Evidence for the Division of Attention Among Non-contiguous Locations. *Visual Cognition*, 5(1/2), 217-256.
- Hopf, J.-M., Luck, S. J., Boelmans, K., Schoenfeld, M. A., Boehler, C. N., Rieger, J., & Heinze, H.-J. (2006). The neural site of attention matches the spatial scale of perception. *The Journal of Neuroscience*, 26(13), 3532-3540.
- Houtkamp, R., & Roelfsema, P. R. (2006). The effect of items in working memory in the deployment of attention and the eyes during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 32(2), 423-442.

Chapter 1

- Jonides, J. & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, 43, 346-354.
- Juola, J. F., Bouwhuis, D. G., Cooper, E. E., & Warner, C. B. (1991). Control of attention around the fovea. *Journal of Experimental Psychology: Human perception and performance*, 17(1), 125-141.
- Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, 22, 751-761.
- Kincade, J. M., Abrams, R. A., Astafiev, S.V., Shulman, G. L., & Corbetta, M. (2005). An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *The Journal of Neuroscience*, 25(18), 4593-4604.
- Kim, Y. H., Gitelman, D. R., Nobre, A. C., Parrish, T. B., LaBar, K. S., & Mesulam, M. M. (1999). The large-scale neural network for spatial attention displays multifunctional overlap but differential asymmetry. *NeuroImage*, 9, 269-277.
- Kramer, A., & Hahn, S. (1995). Splitting the Beam: Distribution of Attention Over Noncontiguous Regions of the Visual Field. *Psychological Science*, 6(6), 381-386.
- Kutas, M., & Dale, A. (1997). Electrical and magnetic readings of mental functions. In: Rugg, M. D. (Eds.). *Cognitive neuroscience*. UK: Biddles Ltd.
- Kusunoki M, Gottlieb J, Goldberg M. E. (2000). The lateral intraparietal area as a salience map: the representation of abrupt onset, stimulus motion, and task relevance. *Vision Research*, 40(10-12), 1459-1468.
- LaBerge, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human perception and performance*, 9, 371-379.
- LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, 96, 101-124.
- Ling, S., & Carrasco, M. (2006). When sustained attention impairs perception. *Nature Neuroscience*, 9(10), 1243-1245.
- Liu, T., Fuller, S. & Carrasco, M. (in press). Attention alters the appearance of motion coherence. *Psychonomic Bulletin & Review*.
- Liu, T., Slotnick, S. D., Serences, J. T., & Yantis, S. (2003). Cortical mechanisms of feature-based attentional control. *Cerebral Cortex*, 13, 1334-1343.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279-281.
- Luo, Y-J., Greenwood, P. M., & Parasuraman, R. (2001). Dynamics of the spatial scale of visual attention revealed by brain event-related potentials. *Cognitive Brain Research*, 12, 371-381.
- Mangun, G. R. & Hillyard, S. A. (1987). The spatial allocation of visual attention as indexed by event-related brain potentials. *Human Factors*, 29(2), 195-211.
- McMains, S. A., & Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron*, 42(4), 677-686.
- McMains, S. A., & Somers, D. C. (2005). Processing efficiency of divided spatial attention mechanisms in human visual cortex. *The Journal of Neuroscience*, 25(41), 9444-9448.
- Mesulam, M M. (1981). A cortical network for directed attention and unilateral neglect. *Annals of neurology*, 10, 309-325.
- Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage*, 11(6), 735-759.
- Montagna, B., & Carrasco, M. (2006). Transient covert attention and the perceived rate of flicker. *Journal of Vision*, 6(9), 955-965.
- Mozer, M. C., & Vecera, S. P. (2005). Object- and space-based attention. In L. Itti, G. Rees, & J. Tsotsos (Eds.), *Neurobiology of attention* (pp. 130-134). New York: Elsevier.

- Müller, N. G., Bartelt, O. A., Donner, T. H., Villringer, A., & Brandt, S. A. (2003). A physiological correlate of the "zoom lens" of visual attention. *The Journal of Neuroscience*, 23(9), 3561-3565.
- Müller, M. M., & Hübner, R. (2002). Can the attentional spotlight be shaped like a doughnut? Evidence from steady state visual evoked potentials. *Psychological Science*, 13, 119-124.
- Müller, N. G., & Kleinschmidt, A. (2003). Dynamic Interaction of Object- and Space-Based Attention in Retinotopic Visual Areas. *The Journal of Neuroscience*, 23(30), 9812-9816.
- Müller, N. G., & Kleinschmidt, A. (2004). The attentional 'spotlight's' penumbra: center-surround modulation in striate cortex. *Neuroreport*, 15(6), 977-980.
- Müller, M. M., Malinowski, P., Gruber, T., & Hillyard, S. A. (2003). Sustained division of the attentional spotlight. *Nature*, 424(6946), 309-312.
- Murphy, T. D. & Eriksen, C. W. (1987). Temporal changes in the distribution of attention in the visual field in response to precues. *Perception & Psychophysics*, 42(6), 576-586.
- O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as units of attentional selection. *Nature*, 401, 584-587.
- Ogawa, S., Lee, T. M., Kay, A. R., Tank, D. W. (1990a). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceeding of the National Academy of Science*, 87(24), 9868-9872.
- Ogawa, S., Lee, T., Nayak, A., & Glynn, P. (1990b). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine*, 14, 68-78.
- Olivers, C. N., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 32(5), 1243-1265.
- Peelen, M. V., Heslenfeld, D. J., & Theeuwes, J. (2004). Endogenous and exogenous attention shifts are mediated by the same large-scale neural network. *NeuroImage*, 22, 822-830.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.
- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In: H. Bouma, & D. Bouwhuis (Eds.), *Attention and Performance X* (pp. 531-556). London: Lawrence Erlbaum.
- Posner, M. I., Snyder, C. R., & Davidson B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology*, 109, 160-174.
- Rauschenberger, R. (2003). Attentional capture by auto- and allocues. *Psychonomic Bulletin & Review*, 10, 814-842.
- Ress, D., Backus, B. T., & Heeger, D. J. (2000). Activity in primary visual cortex predicts performance in a visual detection task. *Nature Neuroscience*, 3(9), 940-945.
- Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron* 26, 703-714.
- Roelfsema, P. R., Lamme, V. A., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, 395(6700), 376-381.
- Rosen, B. R., Buckner, R. L., & Dale, A. M. (1998). Event related fMRI: Past, present, and future. *Proceeding of the National Academy of Science*, 95, 773-80.
- Rosen, A. C., Rao, S. M., Caffarra, P., Scaglioni, A., Bobholz, J. A., Woodley, S. J., Hammeke, T. A., Cunningham, J. M., Prieto, T. E., & Binder, J. R. (1999). Neural basis of endogenous and exogenous spatial orienting. A functional MRI study. *Journal of Cognitive Neuroscience*, 11, 135-152.
- Ruz, M., & Lupiáñez, J. (2002). A review of attentional capture: On its automaticity and sensitivity to endogenous control. *Psicológica*, 23, 283-309.
- Sanders, A. F. (1990). Issues and trends in the debate on discrete vs. continuous processing of information. *Acta Psychologica*, 74, 123-167.
- Scharlau, I. (2004). Evidence for split foci of attention in a priming paradigm. *Perception & Psychophysics*, 66(6), 988-1002.
- Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition*, 80, 1-46.

- Serences, J. T. (2004). A comparison of methods for characterizing the event-related BOLD timeseries in rapid fMRI. *Neuroimage*, 21(4), 1690-1700.
- Serences, J. T., Schwarzbach, J., Golay, X., Courtney, S. M., & Yantis, S. (2004). Control of object-based attention in human cortex. *Cerebral Cortex*, 14, 1346-1357.
- Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychological Science*, 16, 114-122.
- Serences, J. T. & Yantis, S. (2006). Selective visual attention and perceptual coherence. *Trends in Cognitive Sciences*, 10, 38-45.
- Serences, J. T., & Yantis, S. (2007). Spatially selective representations of voluntary and stimulus-driven attentional priority in human occipital, parietal, and frontal cortex. *Cerebral Cortex*, 17(2), 284-293.
- Shaw, M. L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal Experimental Psychology Human Perception and Performance*, 3, 201-211.
- Shomstein, S., & Behrmann, M. (2006). Cortical systems mediating visual attention to both objects and spatial locations. *Proceedings of the National Academy of Sciences USA*, 103(30), 11387-11392.
- Shomstein, S., & Yantis, S. (2004). Control of attention shifts between vision and audition in human cortex. *The Journal of Neuroscience*, 24, 10702-10706.
- Shulman, G. L., Remington, R. W., & MacLean, J. P. (1979). Moving attention through physical space. *Journal of Experimental Psychology: Human perception and performance*, 5, 522-526.
- Shulman, G. L., Sheehy, J. B., & Wilson J. (1986). Gradients of spatial attention. *Acta Psychologica*, 61, 167-181.
- Silver, M. A., Ress, D., & Heeger, D. J. (2005). Topographic maps of visual spatial attention in human parietal cortex. *J Neurophysiol.*, 94(2), 1358-1371.
- Slotnick, S. D., Schwarzbach, J., & Yantis, S. (2003). Attentional inhibition of visual processing in human striate and extrastriate cortex. *Neuroimage*, 19(4), 1602-1611.
- Somers, D. C., Dale, A. M., Seiffert, A. E., & Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences USA*, 96, 1663-1668.
- Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 31(2), 248-261.
- Sperling, G. & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review*, 102(3), 503-532.
- Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, 50, 184-193.
- Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, 51, 599-606.
- Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human perception and performance*, 20, 799-806.
- Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review*, 11, 65-70.
- Thompson, K. G., & Bichot, N. P. (2005). A visual salience map in the primate frontal eye field. *Progress in brain research*, 147, 251-262.
- Tomita, H., Ohbayashi, M., Nakahara, K., Hasegawa, I., & Miyashita, Y. (1999). Topdown signal from prefrontal cortex in executive control of memory retrieval. *Nature*, 401, 699-703.
- Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., & Dale, A. M. (1998). The retinotopy of visual spatial attention. *Neuron*, 21(6), 1409-1422.
- Tsal, Y. (1983). Movements of attention across the visual field. *Journal of Experimental Psychology: Human perception and performance*, 9, 523-530.

- Van der Heijden, A. H. C., Schreuder, R., & Wolters, G. (1985). Enhancing single-item recognition accuracy by cuing spatial locations in vision. *Quarterly Journal of Experimental Psychology*, 37A, 427-434.
- Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, 123, 146-160.
- Woldorff, M. G., Hazlett, C. J., Fichtenholtz, H. M., Weissman, D. H., Dale, A. M., & Song, A. W. (2004). Functional parcellation of attentional control regions of the brain. *Journal of Cognitive Neuroscience*, 16, 149-165.
- Woodman, G. F. & Luck, S. J. (in press). Do the contents of visual working memory automatically influence attention selection during visual search?, *Journal of Experimental Psychology: Human Perception and Performance*.
- Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 601-621.
- Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., & Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial attention shifts *Nature Neuroscience*, 5, 995-1002.
- Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, 396, 72-75.
- Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research*, 39, 293-306.
- Yeshurun, Y., & Carrasco, M. (2000). The locus of attentional effects in texture segmentation. *Nature Neuroscience*, 3, 622-627.

Chapter 2

Remembered but unused:

**the accessory items in working memory that do not
guide attention**

Based on: Peters, J. C., Goebel, R., & Roelfsema, P. R. Remembered but unused:
the accessory items in working memory that do not guide attention. (*submitted*)

Abstract

If we search for an item, a representation of this item in our working memory guides attention to matching items in the visual scene. We can hold multiple items in working memory. Do all these items guide attention in parallel? Are they matched indiscriminately to the items in the display? We asked participants to detect a target object in a stream of objects while they maintained a second item in memory for a subsequent task. On some trials, we presented this memory item as a distractor in the stream. Subjects did not confuse these memory items with the search target, as the false alarm rate on trials where the memory item was presented in the stream was comparable to that on trials with only regular distractors. However, a comparable performance does not exclude that the memory items are processed differently from normal distractors. We therefore recorded the subject's event-related potentials (ERPs) evoked by the search targets, memory items and regular distractors. As expected, the ERPs evoked by search targets differed from those evoked by distractors. Search targets elicited an occipital selection negativity and a frontal selection positivity indexing selective attention, while the P3b component, which is thought to reflect the matching of sensory events to memory representations, was enhanced for targets compared to distractors. Remarkably, the ERPs evoked by memory items were indistinguishable from the ERPs evoked by normal distractors. This implies that only the search target has a special status in working memory that is not shared by the other items. These other, 'accessory' items do not guide attention and are excluded from the matching process.



Introduction

Visual search relies on the interaction between memory and vision. To find an item, we maintain a representation of this item in our working memory, and match this representation to the visual scene (Wolfe, 1994). In recent years, much has been learned about the neurophysiology of working memory. Many neurons that code the visual features of an item when it is perceived also maintain their activity if the item has to be remembered in the absence of the visual input. Persistent neuronal activity as a correlate of working memory has been observed in the frontal cortex (Rainer, Asaad, & Miller, 1998; Courtney, Ungerleider, Keil & Haxby, 1997; Miller, Erickson, & Desimone, 1996; Wilson, O'Scalaidhe & Goldman-Rakic, 1993), but also in many areas of the visual cortex (Corbetta, Kincade, & Shulman, 2002; Supér, Spekreijse & Lamme, 2001; Chafee & Goldman-Rakic, 2000; Chelazzi, Duncan, Miller, & Desimone, 1998; Miller & Desimone, 1994; Chelazzi, Miller, Duncan, & Desimone, 1993; Gnadt & Andersen, 1988; Fuster & Jervey, 1981). During visual search, the neuronal representation of the target in working memory (the 'search template'; Duncan & Humphreys, 1989) has to interact with the incoming visual information in order to identify matching items and to specify their location. Models of visual search usually assume that the neurons representing the search template in higher areas provide a top-down signal to visual areas that enhances the activity of neurons responding to target features in the display (Hamker, 2005; Usher & Niebur, 1996; Desimone & Duncan, 1995; Phaf, Van der Heijden, & Hudson, 1990; Bundesen, 1990; Duncan & Humphreys, 1989). Neurophysiological studies in monkeys have confirmed this idea (Tomita et al., 1999). Neurons in the visual and frontal cortex have a stronger response if their receptive field falls on an item that the monkey is looking for compared to when it falls on a distractor (Bichot, Rossi, & Desimone, 2005; Chelazzi et al., 1998, 1993; Schall & Hanes, 1993). At a psychological level of description, the enhancement of neuronal firing rates is associated with a shift of visual attention to the matching item in the display. In short, working memory representations focus visual attention on matching items in the display (Desimone & Duncan, 1995; Wolfe, 1994; Bundesen, 1990; Phaf, Van der Heijden, & Hudson, 1990; Duncan & Humphreys, 1989).

The total number of items that can be stored in visual working memory is approximately four (Cowan, 2001; Luck & Vogel, 1997). In many cases, some of these items are not used immediately but only become relevant after a delay (Ballard, Hayhoe, & Pelz, 1995). It is an important theoretical question whether all items in working memory have a similar status. If so, the items in working memory that are stored for later use (hereafter called accessory memory items) should also guide attention to matching items in the display, just like the search template. A number of recent studies (Woodman & Luck, in press; Houtkamp & Roelfsema, 2006a; Soto, Heinke, Humphreys, & Blanco, 2005; Downing & Dodds, 2004) investigated this question and found that the search template has a

much stronger influence on the deployment of attention than the accessory memory items.

These results imply that items in visual working memory can attain at least two different states (see also Cowan, 2001). The first is that of the search template that guides attention to the matching items in the display. Accessory items, on the other hand, have less influence on the deployment of attention. Nevertheless, the aforementioned studies do not fully agree on the residual effects of these accessory items. Downing and Dodds (2004) and Houtkamp and Roelfsema (2006a) asked subjects to search for an item, while subjects held another item in memory for a subsequent task. They found that the accessory memory item caused little or no interference with visual search, even if there was a matching item in the display. Their results support a model of working memory where the accessory items are in an isolated state so that they cannot influence the deployment of attention (Figure 1A).

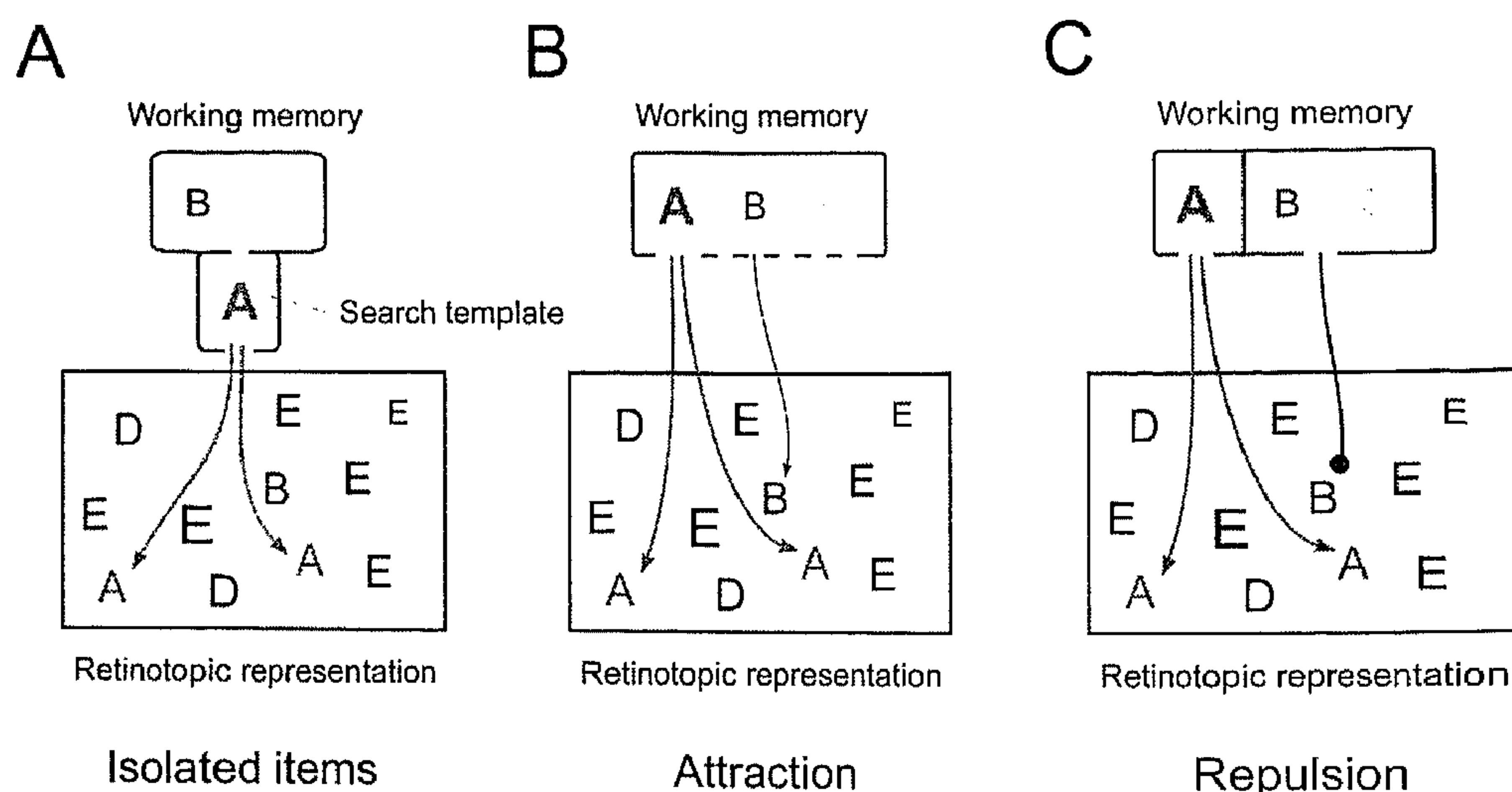


Figure 1. Three models of the organization of items in working memory. All models attribute a special status to the 'search-template': the item in working memory that provides top-down signals to visual areas to focus attention onto matching items in the display. (A) The accessory memory items are isolated from visual processing, as only the search template can influence the deployment of visual attention during search. (B) The accessory memory items attract attention, although to a lesser degree than the search template. (C) The subject strategically uses the accessory memory items as a 'template for rejection'. The accessory memory items send top-down signals to visual areas that repel attention.

Soto and colleagues (2005), however, did observe attraction of attention by items stored for a later task. Their subjects searched for a tilted line among vertical lines in a display where all lines were embedded in colored shapes. In some trials, one of these shapes matched an accessory memory item (in color or shape), and reaction times were shortest if the target line appeared in this item. Likewise, the results of Olivers, Meijer and Theeuwes (2006) revealed a similar

attraction of attention by accessory memory items in an irrelevant singleton paradigm. These results suggest an ‘attraction’ model in which the accessory items exert a top-down influence so that corresponding display items attract attention (Figure 1B). However, another study obtained yet different results. Woodman and Luck (in press) found that visual search was *faster* when the accessory memory item appeared in the search array as a distractor, an effect which also was observed in one of the experiments by Downing and Dodds (2004). Based on this result, Woodman and Luck (in press) suggested that items in working memory other than the search template may even *repel* attention. Participants may strategically avoid attending the distractors that match the memory item in the search array, as if its representation in working memory can act as a “*template for rejection*” (Woodman & Luck, in press) (Figure 1C).

Thus, although the previous studies agree that the accessory memory items have a different representation than the search template, their exact status remains unclear. At first sight, it seems difficult to reconcile the apparently conflicting results. We note, however, that the previous studies focused on reaction times and accuracy: measures that reflect the final outcome of a number of processing steps (e.g., Sanders, 1990). During visual search, display items undergo sensory processing, may or may not attract attention and may or may not be matched against the search template. It is therefore possible that processing of a distractor that matches an accessory memory item differs from the processing of regular distractors at more than one processing step, while these differences are not reflected in the subject’s performance. The accessory memory item may, for example, interfere because it attracts attention, but this effect may be offset because it is also rejected faster by the matching process if it acts as a ‘template for rejection’.

In the present study we have recorded event-related potentials (ERPs) to gain insight in these component processes. Participants searched for a target object in a stream of objects, while they memorized a second object for a subsequent task (Figure 2). We will compare ERPs evoked by the search targets, memory items and regular distractors. Specifically, we will measure the *occipital selection negativity* (OSN; Harter, Aine & Schroeder, 1982; Previc & Harter, 1982) and *frontal selection positivity* (FSP; Anllo-Vento, Luck, & Hillyard, 1998; Kenemans, Kok, & Smulders, 1993) as indices of selective attention, and the later P3b (or P300; hereafter referred to as P3) to probe the matching of stimuli to memory (for a review, see Kok, 2001).

Methods

Participants

Nineteen healthy volunteers (7 males; mean age: 21.8 years) with normal or corrected-to-normal visual acuity were paid for their participation in this study. They were all right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971) and gave their informed consent to participate. All procedures were approved by the ethics committee of the Faculty of Psychology of Maastricht University.

Stimuli and design

The participants were tested in a soundproof and electrically shielded room while seated in front of a 17 inch CRT monitor (60 Hz refresh rate). Stimuli were presented and responses were recorded using the Presentation software package (Neurobehavioral Systems, San Francisco, USA). Trials began with the presentation of a fixation cross (1200 ± 200 ms) in the middle of a grey screen. Subsequently, two randomly chosen shapes were presented (1000 ms each with 150 ms in between, see Figure 2). We used six complex unnatural objects (Blurred Outlined Random Tetris Shapes; Bledowski et al., 2006) with a size of approximately $6.0^\circ \times 6.0^\circ$, presented at a random location within 2.1° of the fixation cross. These complex objects cannot be easily verbalized and are therefore likely to be retained in visual working memory. After 150 ms, a cue (a “1” or a “2” with equal probability) was presented for 1000 ms to indicate which of the two objects served as the search target. The other object had to be held in memory for a subsequent object matching task. We presented both objects before the cue to avoid potential differences in encoding strategies for targets of the search and memory task. After a delay of 1000 ± 200 ms we presented the search stream that consisted of a sequence of three randomly chosen distractors and the ‘stimulus-of-interest’ (i.e., the search target, the memory item or another, regular distractor) that defined the stream type. Thus, every search stream contained four unique stimuli, including one ‘stimulus-of-interest’ presented at a random position in the stream (we ensured that every position was selected equally often across trials). The ‘stimulus-of-interest’ distractor did not differ from the other distractors, but we only used this distractor in our ERP analyses to equate the number of trials in the three conditions. We presented the stimuli for 800 ms with an inter stimulus interval of 200 ms. After the stream, we presented a response cue and the participant indicated whether the target was present (left mouse button) or absent (right mouse button). Streams contained equally often the search target, the memory item or only distractors (31% each). A fourth type of trials were catch trials (7%) in which the search target was presented twice to encourage the participants to continue their search when they detected a target. The participants were instructed to press the scroll button of the mouse on these catch trials.

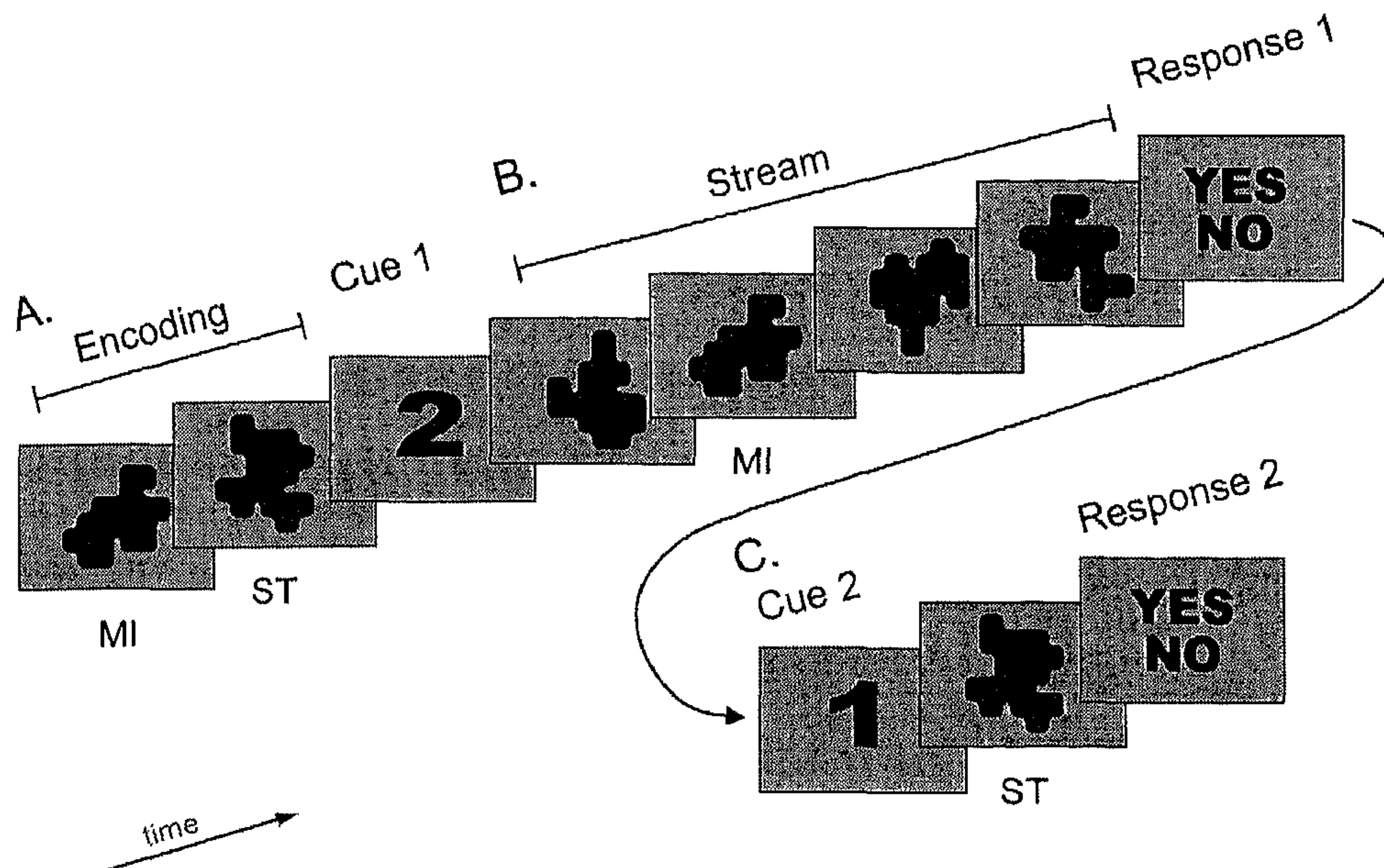


Figure 2. Sequence of events during a trial. (A) In the encoding phase, we presented two objects that were followed by a number indicating which of the two objects was the target (search target; ST) in the subsequent search stream. The other object (memory item; MI) had to be remembered for a subsequent memory task. (B) The search stream was composed of three distractors and a 'stimulus-of-interest' (i.e., the search target, memory item, or a distractor). Then a response cue appeared which instructed the participants to indicate whether the target had been presented in the stream. (C) A second cue (a '1' if the first cue was '2' and vice versa) signaled that in the succeeding task, the memory item would become the new target while the search target could be ignored. We then presented either the memory item, the search target or a distractor as a memory probe. The response cue instructed the participants to indicate whether this probe matched the memory item. Note that in this particular example, subjects should respond 'no' in both tasks.

The subsequent memory task started with a second cue (a '1' if the first cue was '2' and vice versa) signaling that the memory item had now become the target (Figure 2). As memory probe, we presented the memory-item (match) on 33.3% of the trials, the search target of the previous stream (mismatch) on the 33.3% of trials or a distractor (mismatch) on the remaining 33.3% of trials, for a duration of 800 ms. A response cue appeared after an interval of 200 ms, and the participants had to indicate whether the probe matched the memory item (or not) by clicking the left (right) mouse button. All combinations of search and memory task conditions occurred equally often. Participants received auditory feedback after each button response. Subjects heard a high tone (100 ms, 1000 Hz) after a correct response, a low tone (500 Hz) after an error, and a ringing sound when they failed to respond within 5 s. In addition, we gave information about the subject's average RT and error percentage after each run.

Procedure

We instructed the participants to respond as fast and accurately as possible and not to speak, blink, or move their eyes away from the fixation cross during stimulus presentation. Prior to the EEG measurement, participants were trained on a task that was similar to the main task, except that the four stimuli of the search array were presented simultaneously (one object on a jittered position in each quadrant). Participants performed 48 trials of this task and they subsequently practiced the main task (58 trials). After this training session, participants performed four blocks (14 minutes each) of 58 trials, resulting in a total of 232 trials. The entire experiment including the practice session and placement of the electrode cap lasted about 2½ hours.

Electrophysiological recording and data analysis

The electroencephalogram (EEG) was recorded from 29 tin scalp electrodes (extended version of the 10/20 system) with reference electrodes placed at the mastoids. Signals were collected using the left mastoid as reference and re-referenced off-line to the average activity at the mastoids. Horizontal and vertical electrooculograms (EOG) were recorded with bipolar electrodes placed at the external canthi and above and below the left eye. Electrode impedance was kept below 5 kOhm for all electrodes.

We band-pass filtered the EEG data between 0.01 and 30 Hz (filter slope 12dB/octave), sampled at 250 Hz and subtracted the base-line activity in a 100 ms pre-stimulus interval. We analyzed the responses in a time window from -200 to 950 ms relative to stimulus onset. Only trials on which subjects gave correct responses for both tasks were analyzed. Catch trials and trials with artifacts (samples exceeding $\pm 75 \mu\text{V}$) were excluded from subsequent analyses.

Electrode sites were collapsed into a frontal (Fp1, F3, F7, Fz, AFz, Fp2, F4, F8), centroparietal (C3, CP3, P3, Cz, CPz, Pz, C4, CP4, P4), and occipitotemporal (T5, O1, Oz, T6, O2) region of interest (ROI, see Figure 3). The mean amplitude of the FSP (250-450 ms), OSN (250-350 ms), and the P3 (350-600 ms) were evaluated within respectively the frontal, occipitotemporal and centroparietal ROI. Mean amplitudes of the FSP, OSN, and P3 evoked by the search target and the memory item were compared to those evoked by a regular distractor using pairwise, two-tailed t-tests. Two participants were excluded from the analysis, since they had a low accuracy on the secondary memory task (<81%), in combination with a poor fixation during the search task.

Results

The subjects' average accuracy was high in the search task (93.7% correct) as well as in the memory task (90.8%). To investigate whether subjects sometimes confused the accessory memory item with the search target, we compared the trials where the memory items appeared as a distractor in the stream to trials with only regular distractors not in memory. The probability of a false alarm on memory item trials (5.0%) was similar to that on distractor trials (4.1%) ($t(16) = 1.14$; $p > .25$). Thus, the memory items behaved as normal distractors during the search task. Nevertheless, the performance on the memory task was very good, which indicates that the accessory memory item was well remembered during the preceding search task. These performance data, taken together, indicate that the subjects memorized both items, while they were able to keep the memory representation of the search target separate from the representation of the accessory memory item.

We next investigated whether the presence of the search target or memory items in the search stream influenced the performance in the subsequent memory task. The probability of a correct rejection of the search target in the memory task increased from 89.2 % to 96.1% if it had appeared as a target in the preceding search stream ($t(16) = 3.95$; $p < .002$). This suggests that subjects were able to refresh their memory representation of the search target if it appeared in the stream. In contrast, the presence of the accessory memory items in the search stream did not have a strong influence on performance in the memory task ($t(16) = 1.81$; $p > .05$).

ERPs evoked by targets, accessory memory items and distractors

We based our ERP analysis on all trials where the subjects gave a correct response in both tasks (85.5% of the trials). We first investigated whether the amplitude of the ERP components interacted with the position of the critical item in the stream. We did not observe interactions between stimulus type and position in the stream, and we therefore pooled responses across all the positions in the stream for the subsequent analyses. Figure 3 shows the average ERP waveforms evoked by the search targets, the accessory memory items, and the regular distractors that were embedded in the search stream. The early ERP components (P1, N1, and P2) evoked at latencies up to 200 ms can be seen to be similar across all types of items, but at longer latencies the ERPs evoked by the targets differ from those evoked by the distractors. The differences between target and distractor processing are also clearly evident in Figure 4, which displays grand average ERPs at the frontal, centroparietal, and occitotemporal ROI. The target and distractor ERPs start to differ at a latency of approximately 250 ms above the frontal and occitotemporal cortex, and this is followed by a later difference in responses above the centroparietal cortex. Importantly, the ERPs evoked by the memory items appear to be similar to those evoked by the distractors.

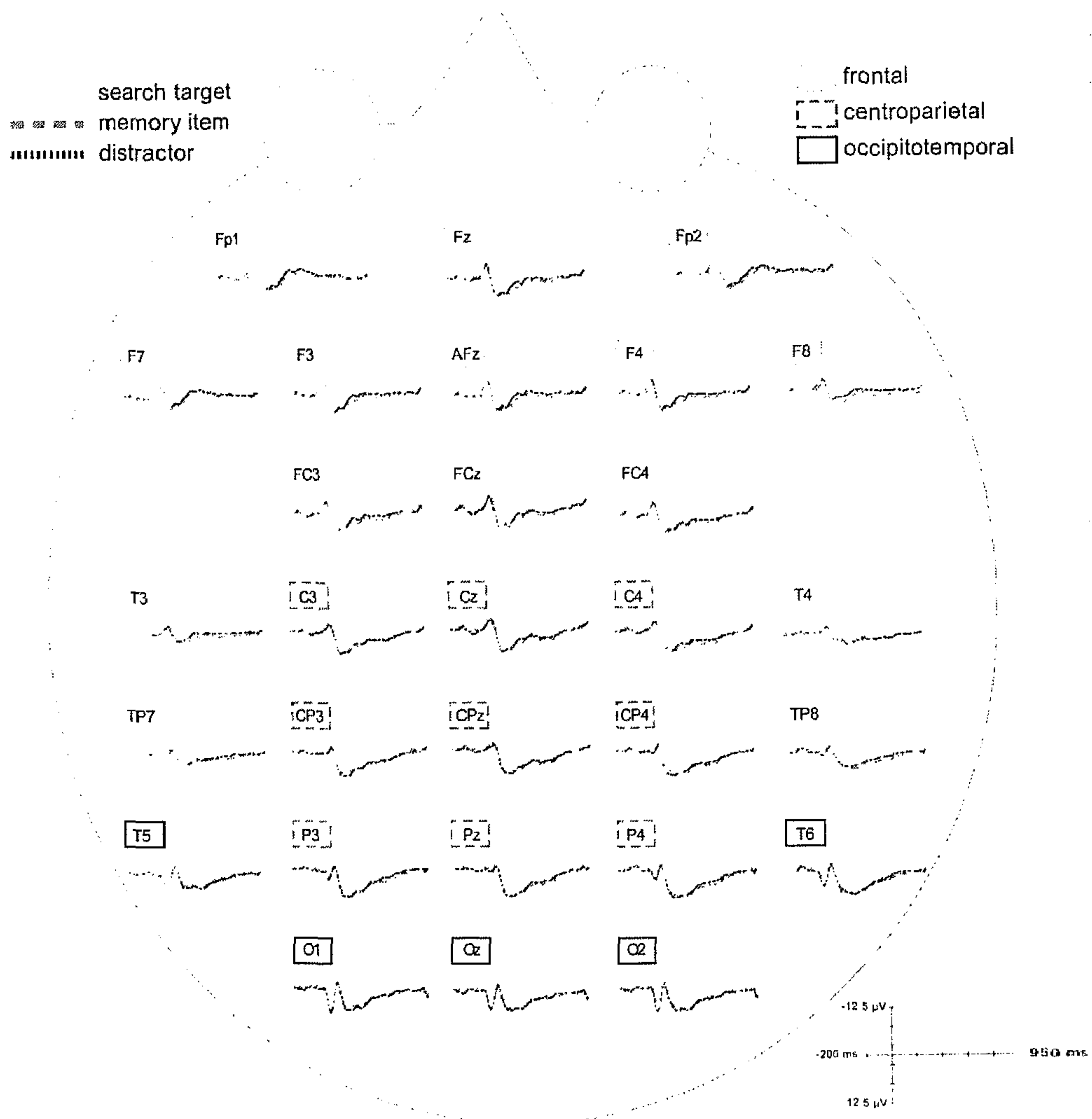
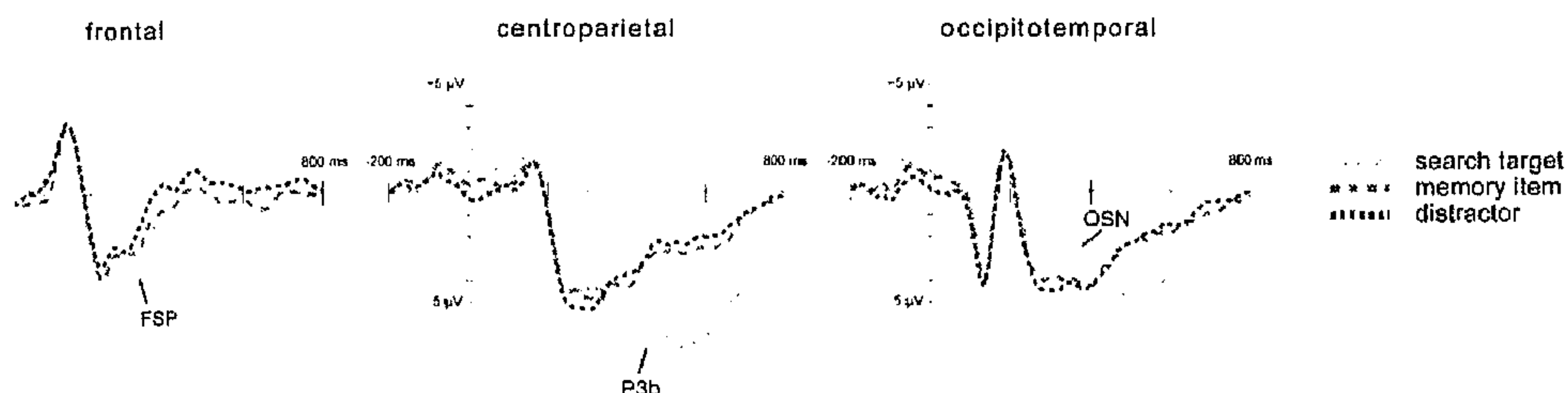


Figure 3. Grand average ERPs for the search target, memory item, and regular distractors presented in the search stream; shown for every scalp electrode site. Frames indicate assignment of the electrodes to one of three Regions-of-Interest (ROIs) for the subsequent analyses.

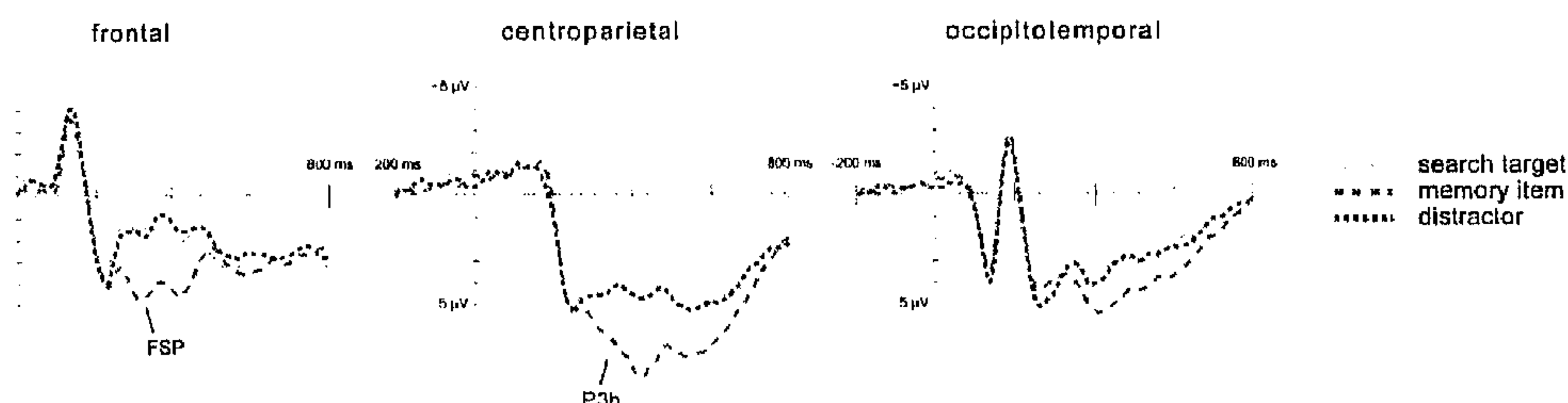
Our statistical analysis supports this description. Planned comparisons confirmed that the search target elicited a more negative response than the distractor at the posterior electrodes (OSN, 250-350 ms: mean difference = 1.71 μ V; $t(16) = 4.92$; $p < .0002$), while it evoked a more positive response than distractors at the anterior electrodes (FSP, 250-450 ms: diff = 1.42 μ V; $t(16) = 2.15$; $p < .05$). In addition, the centroparietal P3 was strongly enhanced for search targets compared to distractors (350-600 ms: diff = 3.30 μ V; $t(16) = 4.44$; $p < .0005$).

Not all items in working memory guide attention: ERP findings



4. Average evoked potentials in the search task recorded from frontal, centroparietal, and occipitotemporal electrodes. Note that the ERP evoked by the accessory memory items is similar to that evoked by regular distractors, while it differs from the ERP evoked by the search target.

In contrast, the response evoked by the memory items was similar to the response evoked by regular distractors. We did not observe significant differences in the amplitude in any of the time intervals (t -tests, all p s > .19). This suggests that the processing of the memory items was similar to that of regular distractors, while distinct from the processing of search targets. Thus, items in the search stream matching the accessory memory item do not appear to attract attention and do not have to be matched against the information in working memory.



5. Evoked potentials in the secondary memory task recorded from frontal, centroparietal, and occipitotemporal electrodes. Now, the memory item ERP differs from the distractor ERP, whereas the search ERP does not.

In the secondary memory task, the roles of the search target and memory item were reversed. Now the memory item was a target while the former search target became a distractor. We expected that the memory items, which were used as distractors during the search task, would now be processed differently. This expectation was confirmed by the data (Figure 5). Now the memory item evoked stronger responses than distractors in the FSP (frontal ROI; diff = 2.71 μ V; $t(16) = 5.23$; $p < .0001$) and P3 (centroparietal ROI; diff = 2.67 μ V; $t(16) = 5.41$; $p < .00006$) latency range, but not in the OSN time-window (occipitotemporal ROI; diff = 0.35 μ V; $t(16) = 0.35$; $p > .7$). However, the response evoked by former search targets resembled the response evoked by the distractors (t -tests, all p s > .6).

Discussion

In many tasks we store items in memory for later use. Here we have investigated if and how these items influence the processing of visual stimuli. The data clearly indicate that the accessory memory items are stored in an inactive state with little influence on visual processing. At a behavioral level, we found that the false alarm rate caused by the memory items was similar to the rate caused by the regular distractors not in memory. Thus, our subjects were well able to memorize two similar items while using only one of them as a template during visual search. These results are in accordance with previous studies that also showed that accessory memory items cause little interference during visual search (Houtkamp & Roelfsema, 2006a; Downing & Dodds, 2004).

In addition, we addressed the possibility that there may exist differences between the processing of memory items and distractors that are not reflected by behavioral measures. Our study is the first to compare the neurophysiological responses evoked by memory items, search targets and distractors. We obtained two neurophysiological results that have a straightforward interpretation. First, we observed a clear signature of the search template, because the ERPs evoked by items in the stream that matched the search targets clearly differed from the ERPs evoked by distractors. Target and distractor ERPs differed during an early phase as well as during a later phase of processing. The frontal (FSP) and occipital (OSN) response differences are related to the selective processing of task-relevant stimuli (e.g., Anllo-Vento, Luck, & Hillyard, 1998), whereas the later P3b component is thought to reflect the matching of stimuli to memory representations (variously interpreted as “context updating” [Donchin & Coles, 1988]; “event categorization” [Kok, 2001], and “template matching” [Chao Nielsen-Bohlman, & Knight, 1995; Squires Hillyard, & Lindsay, 1973]; for a review, see Kok, 2001).

Second, we found that the ERPs evoked by accessory memory items are indistinguishable from the ERPs evoked by distractors not in memory. Our design was sensitive to the difference in processing between targets and distractors, and the absence of a difference between memory items and distractors therefore implies that the accessory memory items do not share those properties with the search template in memory that cause targets in the search stream to be processed differently from distractors. In other words, the memory items do not guide attention and are not matched against the incoming visual information. Our results therefore support a model of working memory organization where the accessory memory items are in an isolated state and do not feed back to visual cortical areas to influence processing (as schematically indicated in Figure 1A). In contrast, the search template appears to occupy a special ‘slot’ or partition in working memory. It is the only item that guides attention and that is matched against the incoming visual information (see also Houtkamp & Roelfsema, 2006a; Downing & Dodds, 2004).

Blockade by the search template

In the introduction we reviewed two studies that reached completely different conclusions regarding the influence of the accessory memory items on visual processing. Soto and coworkers (2005) as well as Olivers et al. (2006) found that display items matching an accessory memory item attracted attention, while Woodman and Luck (in press) observed that subjects used the accessory items in working memory to strategically avoid processing of irrelevant display items. How can we reconcile these findings with the studies that observed little influence of the accessory memory items, including the present one? A critical difference between studies that did and did not observe an influence of the memory items on visual processing is related to the type of search task. The studies that did not observe effects of the accessory memory items selected the search target from the same set of items as the memory items, and changed the identity of the search target from one trial to the next. These search targets therefore had to be stored in working memory at the beginning of every trial. The subjects of the study of Soto et al. (2005), Olivers et al. (2006) and Woodman and Luck (in press), on the other hand, always searched for the same targets. These studies therefore used a so-called ‘consistent mapping’ design (see Shiffrin & Schneider, 1977) where the search target comes from a set of items that is disjoint from the set of potential distractors. After sufficient training, the set of search targets becomes so familiar that it is unnecessary to show the target at the beginning of a trial, and it may therefore cease to occupy storage space in working memory (Woodman, Luck & Schall, 2007). Consistent mapping changes a search task into a categorization task, and Shiffrin & Schneider (1977) showed that subjects become very efficient if they only have to decide whether one of the display items belongs to a fixed target category.

Figure 6 illustrates how the distinction between consistent and varied mapping conditions permits a reconciliation of present and previous results. If the search template occupies space in working memory, then it blocks the influence of the accessory memory items on visual processing so that they are in an isolated state (Figure 6A). However, if the search target remains the same across trials and ceases to occupy space in working memory, this blockade is released (Figure 6B) and now the accessory memory items may interfere with visual search (Soto et al., 2005; Olivers et al., 2006) or may even be used strategically as a template for rejection to make visual search more efficient (Woodman & Luck, in press). Preliminary findings of Oh and Kim (2003) confirm this interpretation. They asked participants to search for either a specific shape that occupied space in working memory or for a vertically symmetric object (regardless of its shape) that did not. In accordance with the model of Figure 6, an accessory memory item only interfered with the visual search when the subjects searched for a symmetric item that was not held in working memory, while interference was not observed when the subjects searched for a specific shape that had to be memorized. The model of Figure 6 receives additional support from a recent study by Woodman, Luck &

Schall (2007). Their subjects performed a visual search for a constant target or for a target that changed from trial to trial. In the meantime, the subjects had to either remember 4 colors or none. This additional memory load interfered with visual search, but only if the search target was varied across trials, in accordance with the idea that constant targets do not occupy space in working memory.

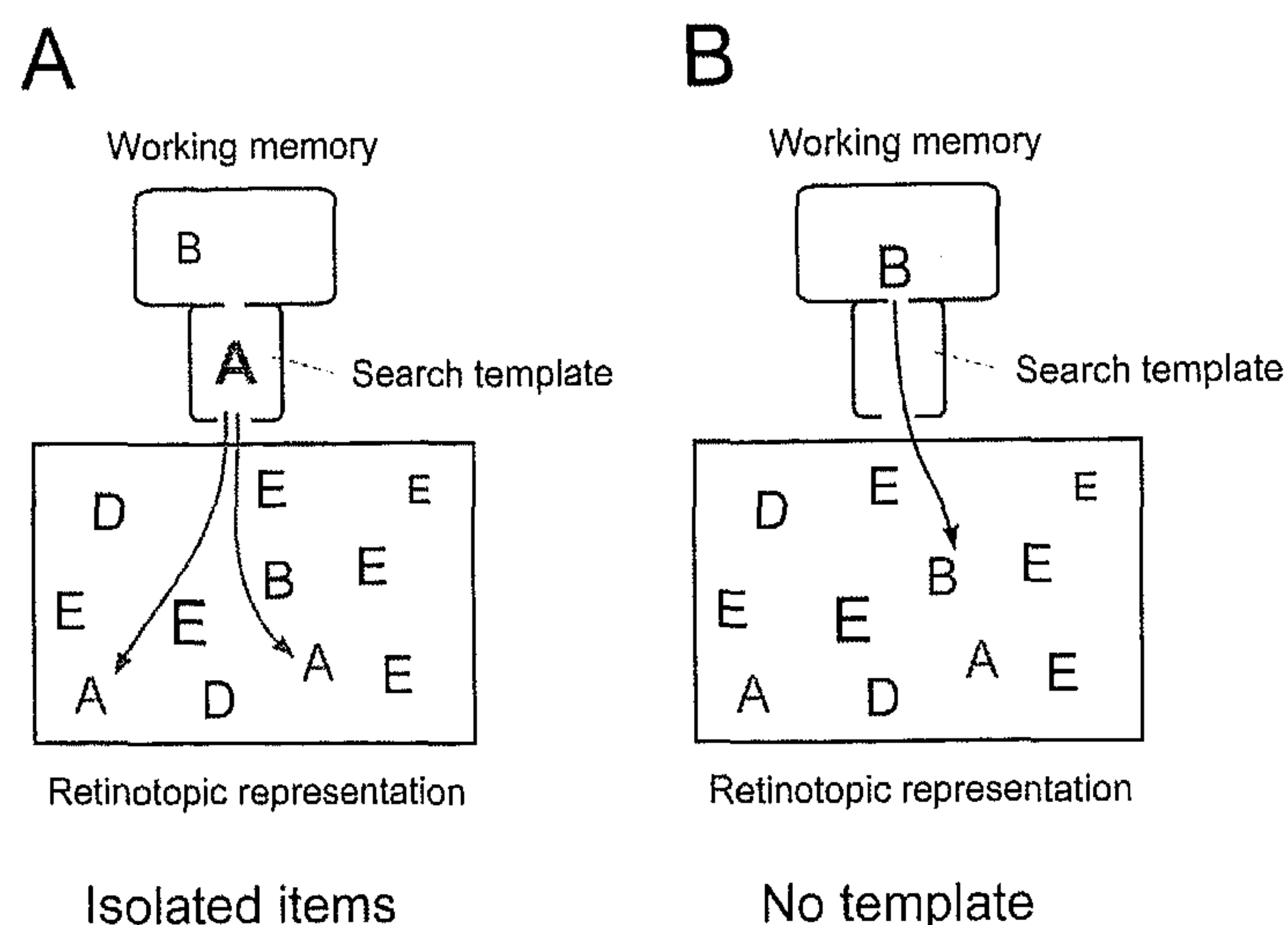


Figure 6. How the search template may influence the effect of accessory memory items. (A) Storage of a search template blocks the influence of the accessory memory items on visual processing, and these items are therefore in an isolated state. (B) In the absence of a search template the accessory memory items may gain a small influence on visual processing.

Different states in working memory

The classical theory of Baddeley (e.g. Baddeley, 2000) proposed that working memory is not a unitary system, but that it consists of several memory stores one of which is visual short-term memory. The present results indicate that there also exists a compartmentalization within visual short-term memory. The capacity of visual short-term memory is about four items (Cowan, 2001; Luck & Vogel, 1997), but all the items residing in working memory are not equivalent. There is a special 'slot' in working memory for the search template: the item in memory that guides attention during visual search and that is matched against the items in the visual display. It is not yet clear whether more than one item in working memory can act as a search template. In the present study it was a good strategy for the subjects to maintain the accessory memory items in an inactive state, because if a matching item appeared in the stream then it was a distractor. In a recent study, Houtkamp and Roelfsema (2006b) asked subjects to look for multiple target items in a stream of rapidly presented stimuli. The subjects of this study were unable to use more than a single template at a time, which suggests that the search template is a unique location in working memory that can hold at most a single item. This

interpretation is consistent with the results of a seminal study of Sternberg (1966), who investigated how much time it takes for subjects to match a single character in a visual display to a varying number of characters in memory. Sternberg found that the reaction time of the subjects increased linearly with the number of memory items, as if the characters in memory were serially compared to the visually presented character. In the model of Figure 6, such a serial comparison entails a successive switching of items between the accessory memory buffer and the search template.

To gain insight in the neuronal representation of the search template, it is useful to briefly review the ‘biased competition’ model of Desimone and Duncan (Desimone & Duncan, 1995; see also Miller & Cohen 2001) who proposed that visually presented items engage in a competition in visual cortical areas. This model holds that the search template stored in the frontal cortex provides a top-down biasing signal to the lower visual areas to enhance the representation of visual items matching the search template, in order to increase the probability that these items win the competition and are selected by visual attention. This model has been supported by a number of neurophysiological studies in macaque monkeys. First, it has been shown that neurons in the frontal cortex of monkeys encode the search template (e.g., Miller, Erickson, & Desimone, 1996). These cells are activated when the search target is presented, and they maintain their activity during memory episodes when the search target can no longer be seen. According to the theory, these neurons would act as a source for a top-down biasing signal to lower visual areas. Other studies recorded from neurons in areas of the visual cortex during visual search (Bichot et al., 2005; Chelazzi et al., 1998, 1993), and demonstrated that neurons activated by visual items that match the search template indeed enhance their activity.

These results, taken together, support the idea that the search template is represented by the persistent activity of neurons in the frontal cortex. But how are the accessory memory items maintained? A recent EEG study by Vogel and Machizawa (2004) demonstrated that the storage of multiple items in working memory is associated with extra persistent neuronal activity. A higher memory load was associated with an increase in negativity over the contralateral occipital and parietal cortex. Correspondingly, fMRI studies have demonstrated that the hemodynamic response in frontal and parietal areas increases with memory load (e.g., Cohen et al. 1997). Only few studies have investigated the activity of single neurons in monkeys during tasks where multiple objects have to be memorized. Warden and Miller (2007) showed that the persistent activity of neurons in the prefrontal cortex during a memory delay can code the identity of two objects. A surprising result of this study was that the representation of the first object in memory changed drastically when a second object was loaded. These results, taken together, suggest that the accessory memory items are also stored by the persistent firing of neurons in a distributed memory network. The representation of the accessory memory items therefore resembles the representation of the search

template, and it will be important to investigate if there are differences between the topologies of the memory networks for the search template and the accessory items that can explain why only the former is matched to the incoming stimuli.

A further topic for future research is to understand what happens when the accessory memory item becomes the search target. If an accessory memory item is coded by neurons that do not generate feedback, then changing its status to that of a search template should cause activation of other neurons that do feed back to lower visual areas. Previous studies suggested that the frontal cortex plays an important role in such a status change. Rao, Rainer, and Miller (1997), for example, asked monkeys to first remember the identity of a search target and then presented a search display. Next, the animals had to memorize the location where the target item had been found. Many neurons in the frontal cortex coded the search template representing the target identity in the first part of the trial and subsequently coded the location of the matching visual item, in accordance with the changing task demands. This implies that some neurons in the frontal cortex code those attributes that are of current relevance.

The hypothesis that the frontal cortex is important for switching between accessory memory items and the search template is supported by a recent study by Soto, Humphreys and Heinke (2006). These authors investigated the effect of accessory memory items on visual search in a group of patients with lesions in the frontal lobe. The patients searched for a tilted line among vertical distractors (i.e. the search template itself presumably needed little storage space) while they kept an item in working memory for a subsequent task. Interestingly, the frontal lobe patients suffered from a stronger interference of the memory items than control subjects. This suggests that the frontal cortex play an important role in the storage of the accessory memory items in a format that prevents interference.

We conclude that the present study, together with other recent studies, provide a new understanding of the substructure of visual working memory. Future investigations can now start to address important questions about the neuronal mechanisms for the storage of the search template and the accessory items, and for the different impact that these memory items have on ongoing information processing in areas of the visual cortex.

Acknowledgments

We thank Nikolaus Kriegeskorte for designing the stimuli and Irene Nagel for assistance in some of the measurements. In addition, we gratefully acknowledge the useful comments on earlier versions of the manuscript by Bernadette Jansma and Joel Reithler. This research was supported by NWO-grant No. 402-01-632 to JCP and RG.

References

- Anllo-Vento, L., Luck, S.J., & Hillyard, S. A. (1998). Spatiotemporal dynamics of attention to color: evidence from human electrophysiology. *Human Brain Mapping*, 6, 216–238.
- Baddeley A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4, 417–423.
- Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, 7, 66–80.
- Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. *Science*, 308, 529–534.
- Bledowski, C., Cohen Kadosh, K., Wibrall, M., Rahm, B., Bittner, R. A., Hoechstetter, K., Scherg, M., Maurer, K., Goebel, R., & Linden, D. E. (2006). Mental chronometry of working memory retrieval: a combined functional magnetic resonance imaging and event-related potentials approach. *Journal of Neuroscience*, 26, 821–829.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, 97, 523–547.
- Chafec, M. V., & Goldman-Rakic, P. S. (2000). Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades. *Journal of Neurophysiology*, 83, 1550–1566.
- Chao, L., Nielsen-Bohlman, L. C., & Knight, R. T. (1995). Auditory event-related potentials dissociate early and late memory processes. *Electroencephalography and Clinical Neurophysiology*, 96, 157–168.
- Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, 80, 2918–2940.
- Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, 363, 345–347.
- Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, 386, 604–608.
- Corbetta, M., Kincade, J. M., & Shulman, G. L. (2002). Neural systems for visual orienting and their relationships to spatial working memory. *Journal of Cognitive Neuroscience*, 14, 508–523.
- Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature*, 386, 608–611.
- Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–114.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.
- Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, 11, 357–427.
- Downing, P. E., & Dodds, C. M. (2004). Competition in visual working memory for control of search. *Visual Cognition*, 11, 689–703.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458.
- Fuster, J. M., & Jervey, J. P. (1981). Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science*, 212, 952–955.
- Gnadt, J. W., & Andersen, R. A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental Brain Research*, 70, 216–220.
- Hamker, F. H. (2005). The reentry hypothesis: the putative interaction of the frontal eye field, ventrolateral prefrontal cortex, and areas V4, IT for attention and eye movement. *Cerebral Cortex*, 15, 431–447.

- Harter, M. R., Aine, C., & Schroeder, C. (1982). Hemispheric differences in the neural processing of stimulus location and type: effects of selective attention on visual evoked potentials. *Neuropsychologia*, 20, 421-438.
- Houtkamp, R., & Roelfsema, P. R. (2006a). The effect of items in working memory in the deployment of attention and the eyes during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 423-442.
- Houtkamp, R., & Roelfsema, P. R. (2006b). Only one search template at a time. *Perception*, 35, ECVF Abstract Supplement.
- Kenemans, J. L., Kok, A., & Smulders, F. T. (1993). Event-related potentials to conjunctions of spatial frequency and orientation as a function of stimulus parameters and response requirements. *Electroencephalography and Clinical Neurophysiology*, 88, 51-63.
- Kok, A. (2001). On the utility of P3b amplitude as a measure of processing capacity. *Psychophysiology*, 38, 557-577.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279-281.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal function. *Annual Review of Neuroscience*, 24, 167-202.
- Miller, E. K., & Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science*, 263, 520-522.
- Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, 16, 5154-5167.
- Oh, S. H., & Kim, M. S. (2003). The guidance effect of working memory load on visual search. *Journal of Vision*, 3, 629a.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97-113.
- Olivers, C. N. L., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1243-1265.
- Phaf, R. H., Van der Heijden, A. H., & Hudson, P. T. (1990). SLAM: a connectionist model for attention in visual selection tasks. *Cognitive Psychology*, 22, 273-341.
- Previc, F. H., & Harter, M. R. (1982). Electrophysiological and behavioral indicants of selective attention to multifeature gratings. *Perception & Psychophysics*, 32, 465-472.
- Rainer, G., Asaad, W. F., & Miller, E. K. (1998). Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature*, 393, 577-579.
- Rao, S. C., Rainer, G., & Miller, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, 276, 821-824.
- Sanders, A. F. (1990). Issues and trends in the debate on discrete vs. continuous processing of information. *Acta Psychologica*, 74, 123-167.
- Schall, J. D., & Hanes, D. P. (1993). Neural basis of saccade target selection in frontal eye field during visual search. *Nature*, 366, 467-469.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84, 127-190.
- Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 248-261.
- Soto, D., Humphreys, G. W., & Heinke, D. (2006). Dividing the mind: The necessary role of the frontal lobes in separating memory from search. *Neuropsychologia*, 44, 1282-1289.
- Squires, K. C., Hillyard, S. A., & Lindsay, P. H. (1973). Vertex potentials evoked during auditory signal detection: relation to decision criteria. *Perception & Psychophysics*, 14, 265-272.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, 153, 652-654.
- Super, H., Spekrijse, H., & Lamme, V. A. (2001). A neural correlate of working memory in the monkey primary visual cortex. *Science*, 293, 120-124.

- Tomita, H., Ohbayashi, M., Nakahara, K., Hasegawa, I., & Miyashita, Y. (1999). Topdown signal from prefrontal cortex in executive control of memory retrieval. *Nature*, 401, 699–703.
- Usher, M. & Niebur, E. (1996). A neural model for parallel, expectation-driven attention for objects. *Journal of Cognitive Neuroscience*, 8, 305-321.
- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428, 748-751.
- Warden, M. R. & Miller, E. K. (2007). The representation of multiple objects in prefrontal neuronal delay activity. *Cerebral Cortex*, 17, i41-i50.
- Wilson, F. A., O'Scalaidhe, S. P., Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science*, 260, 1955-1958.
- Wolfe, J. M. (1994) Guided Search 2.0: A Revised Model of Visual Search. *Psychonomic Bulletin & Review*, 1, 202-238.
- Woodman, G. F., Luck, S. J., & Schall, J. D. (2007). The role of working memory representations in the control of attention. *Cerebral Cortex*, 17, i118-i124.
- Woodman, G. F. & Luck, S. J. (in press). Do the contents of visual working memory automatically influence attention selection during visual search? *Journal of Experimental Psychology: Human Perception and Performance*.

Chapter 3

Neural correlates of attentional guidance by items in visual working memory

Based on: Peters, J. C., Roelfsema, P. R., & Goebel, R. Neural correlates of attentional guidance by items in visual working memory. (*in preparation*)

Abstract

If we search for an item, a representation of this item in our working memory (the ‘search template’; Duncan & Humphreys, 1989) guides attention to matching items in the visual scene. Because multiple items can be held in working memory, an important question is whether during visual search all these items guide attention to matching stimuli or only those that are relevant for the task at hand. We used rapid event-related functional magnetic resonance imaging to compare the attentional guiding capabilities of the task-relevant search template with a task-irrelevant item stored in memory for later use. Participants had to detect a face or house target in a stream of superimposed face and house images while they maintained a second face or house in memory (the ‘memory item’) for a subsequent search task. On some trials, we presented this memory item as distractor in the stream of the first search task. Results showed that in the first task the search template signaled category- as well as item-specific information to higher visual areas, as respectively suggested by (object)category-specific sustained increases in activity throughout the search, and transient ‘match enhancements’ (Miller & Desimone, 1994) when the specific target item was encountered. Importantly, the task-irrelevant memory item did not induce such category- or item-specific enhancements, indicating that only the search template guides attention during visual search.

Introduction

In almost all visual attention models, working memory (WM) plays an influential role in the guidance of attention (e.g., Bundesen, 1990; Desimone & Duncan, 1995; Duncan & Humphreys, 1989). For example, the biased competition model (Desimone & Duncan, 1995) proposes that the competition between incoming visual signals for attentive processing can be biased by WM content. Single-cell recordings suggested that prefrontal structures involved in maintaining visual information can indeed provide top-down feedback to lower areas (e.g., Fuster, Bauer, & Jervey, 1985; Miller, Erickson, & Desimone, 1996; Tomita, Ohbayashi, Nakahara, Hasegawa, & Miyashita, 1999). During visual search for instance, feedback from the target representation in WM (further referred to as the search template; Duncan & Humphreys, 1989) seems to enhance activation of neurons in visual areas that are tuned to process corresponding information (Chelazzi, Duncan, Miller, & Desimone, 1998; Chelazzi, Miller, Duncan, & Desimone, 1993), resulting in a competitive advantage for sensory inputs matching the search template to be selected for further attentional processing. Studies by Miller and colleagues (Miller & Desimone, 1994; Miller, Erickson, & Desimone, 1996) have shown that this feedback from WM remains available across distracting inputs, so that it can guide the deployment of attention throughout visual search until the target is found. Additionally, these studies showed that processing of input matching the feedback signal is enhanced compared to non-matching input, an effect which they referred to as “match enhancement” (Miller & Desimone, 1994; p. 521).

Although these results clearly indicate that the search template can guide the deployment of attention, it is not clear whether this is also true for items in WM that are stored for later use (hereafter called accessory memory items). Several recent behavioral studies (Downing & Dodds, 2004; Houtkamp & Roelfsema, 2006; Olivers, Meijer, & Theeuwes, 2006; Soto, Heinke, Humphreys, & Blanco, 2005; Woodman & Luck, in press) investigated this issue using a paradigm in which a target had to be searched in a search array, while another item had to be held in WM for a subsequent task. In some of the trials, this memory item appeared as one of the distractors in the first task. The crucial question was whether more attention would be deployed to this item compared to other distractor items, thereby decreasing search efficiency. The conducted studies did not fully agree on the influence of the accessory memory item on the deployment of attention. In short (see Chapter 2 for a more extensive discussion), Downing and Dodds (2004) and Houtkamp and Roelfsema (2006) observed little or no interference of the memory item with visual search, whereas memory items presented in the search array did seem to attract attention in the studies of Olivers et al. (2006) and Soto et al. (2005). Interestingly, Woodman and Luck (in press) also observed an influence of the memory item on visual search, but in their case search was *faster* when the memory item appeared as a distractor in the search

array. This latter result suggested that subjects could have strategically avoided to attend the memory items in the display, as if its representation in working memory can act as a “template for rejection” (Woodman & Luck, in press).

Since behavioral performance only reflects the outcome of information processing which inherently encompasses several stages (e.g., Sanders, 1990), we performed an Event Related Potential (ERP) study to examine the influence of WM content on the different stages of stimulus processing (see Chapter 2). Participants had to detect a target object in a stream of objects, while maintaining a second object in WM to function as target in a subsequent task. In one third of the trials, this memory item was presented as distractor in the first task. Search targets elicited an occipital selection negativity and a frontal selection positivity indexing selective attention, while the P3b component, which is thought to reflect the matching of sensory events to memory representations, was enhanced for targets compared to distractors. Conversely, the ERPs evoked by memory items were indistinguishable from the ERPs evoked by normal distractors, indicating that memory items did not receive more attention, or were in any other way differently processed than regular distractors.

In the present study, we used functional magnetic resonance imaging (fMRI) to extend our ERP results by investigating the neural mechanisms underlying attentional guidance by WM representations in more detail. Participants searched for a house or face stimulus (*search target*) in a Rapid Serial Visual Presentation (RSVP) of images with superimposed, semi-transparent houses and faces (further referred to as a *search stream* consisting of *search items*), while simultaneously maintaining a second house or face (*memory item*) in WM for a second search task (see Figure 2). Our main question was whether this accessory memory item has a similar top-down influence on visual search processes as the search template. By analysing the attentional modulations in the ‘parahippocampal place area’ (PPA; Epstein & Kanwisher, 1998) and the ‘fusiform face area’ (FFA; Kanwisher, McDermott, & Chun, 1997]), which are respectively specialized in house and face processing, we could specifically study whether the items in WM differentially biased attention towards the house or face attribute of the search items.

Based on previous findings, we had clear expectations regarding the effect of the search template on visual search activity. Firstly, we expected that feedback from the search template would result in sustained enhanced activation in the visual areas specialized for processing the object category to which the search target belonged. That is, if the target was a face, enhanced processing of the face attribute of the search items was expected throughout the entire visual search, resulting in sustained increases in FFA activity. Secondly, when the target would be detected in the search stream, the selection for further attentive processing of the face over the house attribute of the search item would lead to transient “match enhanced” activity in the FFA. However, the present study mainly focused on the influence of the accessory memory item on visual search processes. Based on our

ERP results, no attentional guidance by the memory item was expected during the first search stream. If this accessory memory item would not provide feedback to visual areas, we would expect that the category-specific sustained enhancements in higher visual areas throughout visual search are not influenced by the category to which the memory item belongs. Moreover, no “match enhancements” would be expected in prefrontal or visual areas when the memory item would be encountered in the first search stream.

Methods

Participants

Nine healthy volunteers (4 males; mean age 26.3 years; standard error: 0.9 years) with normal or corrected-to-normal visual acuity were paid for participation in this study. All participants were right-handed and gave their informed consent to participate in the study, which was approved by the local Ethical Committee. Two participants were excluded from analyses because of an extremely high percentage of misses in their behavioral responses ($> 35\%$).

Design

Stimuli

Eighteen grayscale photographs of houses and eighteen grayscale front-view photographs of Caucasian male ($n = 9$) and female ($n = 9$) faces served as stimuli (see Serences, Schwarzbach, Golay, Courtney, & Yantis, 2004). The photographs were equal in luminance and unfamiliar to the participant.



Figure 1. Example of a semi-transparently superimposed face/house stimulus. See text and Figure 2 for further details.

Each individual face photograph was semi-transparently superimposed on each individual house photograph (Figure 1), resulting in 324 superimposed face-house images that could be presented in the search streams. House and face stimuli were superimposed rather than presented in two separate search streams to avoid potential confounding effects of eye-movements or spatial attention shifts between items. Moreover, since house and face stimuli were presented at the same location, memory items of another category than the search target had a maximal opportunity to interfere. Note that the relative weighting between the face (75%)

and the house (25%) photograph (i.e., houses were three times as transparent as faces) was based on a behavioral pilot experiment ($n = 10$), to equalize task difficulty between the house detection and face detection task.

Procedure

As illustrated in Figure 2, each trial began with the serial presentation of two randomly chosen faces, houses or a face and a house stimulus (for a duration of 1900 ms each with 100 ms in between). After 100 ms, a cue (a “1” or “2” with equal probability) was presented for 2 s to indicate which of the two stimuli had to be searched for in the following RSVP of search items (i.e., which of the two was the search target). However, participants were instructed to remember the other stimulus (the memory item) for a subsequent search task.

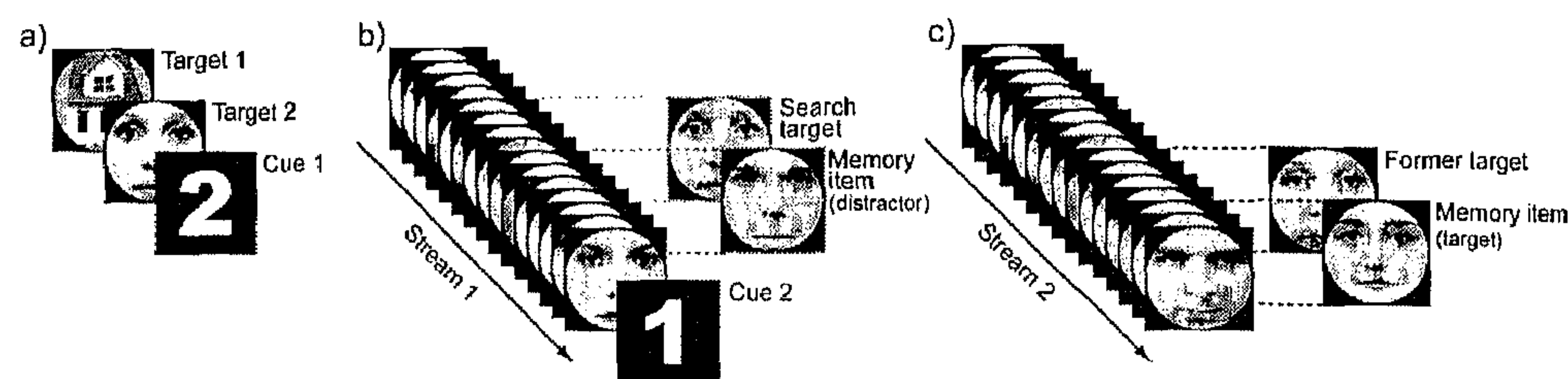


Figure 2. Schematic illustration of the trial design. Each trial consisted of three phases: a). presentation of the two targets and a cue indicating which target (in this example, target 2) had to be searched in b). the first search stream. After the first stream, a second cue signaled that the memory item (target 1) subsequently had to be searched in c). the second search stream. The search/former target (target 2) and the memory item (target 1) could occur in both streams. Note that the target (search target) of the first stream is a distractor (former target) in the second search stream. Conversely, the target (memory item) in the second stream is a distractor (memory item) in the first stream. See text for further details.

We presented both objects before the cue to avoid potential differences in encoding strategies for targets of the first and second task. After a variable cue-delay ($4\text{ s} \pm 2\text{ s}$), the first RSVP of 15 search items (500 ms each; 7500 ms in total) was presented. According to the category of the target (i.e., face or house stimulus) searched in the stream, streams will be referred to as *face search stream* or *house search stream*, respectively. For the face as well as for the house search stream, four types of search stream existed. These four types were defined by two of the fifteen search items referred to as the ‘items-of-interest’ (IOIs), indicating that only these items were of interest for subsequent fMRI analyses. The first search stream could contain the following combination of IOIs: 1.) the search target and the memory item, 2.) the search target and a distractor 3.) the memory item and a distractor 4.) two distractors. Likewise, the following combination of IOIs could occur in the second search stream: 1.) the memory item and former target, 2.) the memory item and a distractor 3.) the former target and a distractor 4.) two distractors. Note that the search target was the target in the first stream, but a distractor (former target) in the second search stream. Conversely, the memory

item served as distractor in the first and as target in the second search stream. Two IOIs could never be presented simultaneously. The remaining 13 items in the search stream functioned as distractors and did not contain representations of the search target or memory item. Although these distractors did not differ from the IOI distractor, only the IOI distractor was further analyzed to keep the number of observations equal for each type of IOI. The individual house and face images were never repeated within a stream. To optimize deconvolution analyses, the onset of the IOI presentation was synchronized to the onset of a volume measurement, with the constraint that the IOI never occurred as the first stimulus of the search stream. The interval between IOIs presented in a search stream was jittered (2 or 4 s). Participants were instructed to search the target throughout the entire stream and indicate its detection with a button response (right index finger). After presentation of the first search stream, a second cue (a '1' if the first cue was '2' and vv.; 2 s duration; presented 500 ms after the last item of the stream) signaled that in the subsequent search stream, the memory item would become the new target while the target of the first search stream (i.e., the former target) should be ignored. Following a variable cue-delay ($4 \text{ s} \pm 2 \text{ s}$), a second search stream was presented having similar characteristics as the first search stream. A small fixation cross was presented in the interval between trials ($6.5 \text{ s} \pm 2 \text{ s}$). All stimuli were presented in the middle of a black screen.

Each of the four types of trials (i.e., two face detection tasks, two house detection tasks, first a face and then a house detection task, or vv.) was presented an equal number of times. Also the four types of search stream occurred equally often and were counterbalanced between the first and the second search stream. Order of trials, order of cues, jittering of cue-delay and intertrial intervals, selection of distractor stimuli in the search stream and the position of the IOIs in the search stream all followed a randomization scheme in which all possibilities occurred equally often across trials.

Image acquisition

Echo-planar images (T2*-weighted; 64x64 imaging matrix, 28 slices, voxel size: $3.5 \times 3.5 \times 3.5 \text{ mm}^3$, no gap, TR/TE = 2 s/35 ms, FA = 90°) covering almost the whole brain were collected on a 3-T Siemens Trio Scanner (Siemens Medical Systems, Erlangen, Germany) using a standard head coil. Functional data were aligned to a T1-weighted high-resolution anatomical image (Magnetization-Prepared Rapid Acquisition Gradient Echo (MPRAGE) sequence; TR/TE = 2.3 s/3.93 ms, FA = 8°). Subjects viewed the stimuli, projected onto a frosted screen using a liquid crystal display projector (VPL-PX21, Sony, Tokyo, Japan), via a mirror mounted to the head coil. Stimuli were presented and responses were recorded using the Presentation software package (Neurobehavioral Systems, San Francisco, USA). Stimulus presentation was synchronized with MR data acquisition.

Each participant performed two runs of the main experiment (1230 volumes in total). A functional localizer of house and face preferring brain regions (160 volumes) was included in the scanning session, using a standard design in which blocks of rapid serially presented face photographs (3 blocks) and house photographs (3 blocks) were alternated with fixation blocks. In addition, each participant performed two additional tasks related to the main experiment. However, the results of these tasks will not be discussed in the present paper for brevity's sake. The tasks were divided over two scanning sessions on the same day (except for one subject, which had the two sessions on different days), with the order of the tasks counterbalanced over subjects. Prior to the fMRI measurement, participants were familiarized with the stimuli and with the different tasks. In all tasks, subjects were instructed to fixate and to respond as fast and accurately as possible.

Data analysis

Behavioral data

Responses occurring after 2 s or more with respect to the target onset were counted as misses. In addition, only responses that occurred within the response time window defined as the interval from two standard deviations below to two standard deviations above the mean reaction time of each participant were included in subsequent analyses. The same individually established response interval was also used for categorizing responses to memory items or former targets and (the pre-specified) distractors as false alarms. Reaction times of correct responses and error percentages were submitted to separate two-way repeated measures analyses of variance (ANOVAs) with stream number (first, second search stream) and category (items in face search stream, items in house search stream) as factors.

fMRI data

Preprocessing of the individual datasets included slice scan time correction, linear trend removal, temporal high-pass filtering, 3D motion correction, transformation into Talairach space (Talairach & Tournoux, 1988), and cortex reconstruction, inflation and flattening as implemented in the BrainVoyager QX software package (version 1.7.9; Brain Innovation, Maastricht, the Netherlands). The first two volumes of each run were discarded to remove T1 saturation effects. No spatial smoothing was applied to the functional data, which were interpolated to a 3 x 3 x 3 mm³ voxel target resolution.

Two types of analyses were performed to study sustained and transient attentional effects during target search. First, a whole brain voxel-wise multiple regression analyses was performed to investigate sustained attentional modulation during the search stream. The design matrix included 8 predictors representing the four main trial phases (i.e., target encoding, cue periods, and target search in

search stream 1 and 2) for the face and house detection task. A ninth predictor was additionally included to separately model the remainder of the stream period whenever a behavioral response was recorded, because participants might have stopped searching after detecting a target. A variant of this design matrix, in which the first house and face search streams were subdivided according to the object category of the memory item, was used to study the influence of the memory item on sustained responses. In addition, a second variant of the design matrix, optimized for studying processes related to cue presentation (i.e., target encoding and search stream predictors were combined for the face and house tasks, but the cue and cuedelay were modeled with separate predictors for conditions and for search streams), was used to study the network involved in updating search template information in WM. In all design matrices, each predictor's boxcar function was convolved with a theoretical Two Gamma hemodynamic response function.

The second analysis was performed to investigate transient attentional effects elicited by IOIs presented in the first search stream. Neural responses to 12 different conditions were estimated. Two conditions represented the target encoding and the cue period (face and house detection trials were combined). IOIs in the first search stream, to which the participant gave a correct response, were represented by one of the eight 'IOI conditions'. These eight conditions were included to model the search target, (regular) distractor, memory item of the same and memory item of a different category (than the search target) of the face and house search streams separately. IOIs to which an incorrect response was given by the participant, were not included in an IOI predictor. Likewise, IOIs presented after participant's response to a search stream stimulus were also not included in one of the IOI conditions, to ensure that the potential attentional capture of memory items was not obscured by including responses to memory items that were possibly no longer attended. The eight corresponding 'IOI conditions' of the second search stream were jointly modeled to minimize the number of predictors. In addition, search stream periods that were not included in one of the IOI conditions were jointly modeled in a 'non-IOI' predictor. Only IOI conditions of the first search stream were further analyzed.

A whole brain multiple regression analysis was performed to investigate which brain areas were sensitive to transient attentional modulations by IOIs in the search stream. To this end, each condition was modeled with a boxcar predictor convolved with a theoretical Two Gamma hemodynamic response function. In addition, multi-subject deconvolution analyses (Glover, 1999; Serences, 2004) were performed on neural responses in the FFA and PPA, to study transient enhancements of IOIs in face and house search streams in detail. For this purpose, each condition was incorporated in the design matrix with 6 delta-function predictors, modeling each of the six time points of the elicited hemodynamic response independently. In order to optimize the estimation of responses to the IOIs, the 'non-IOI' condition was used as baseline instead of the intertrial interval.

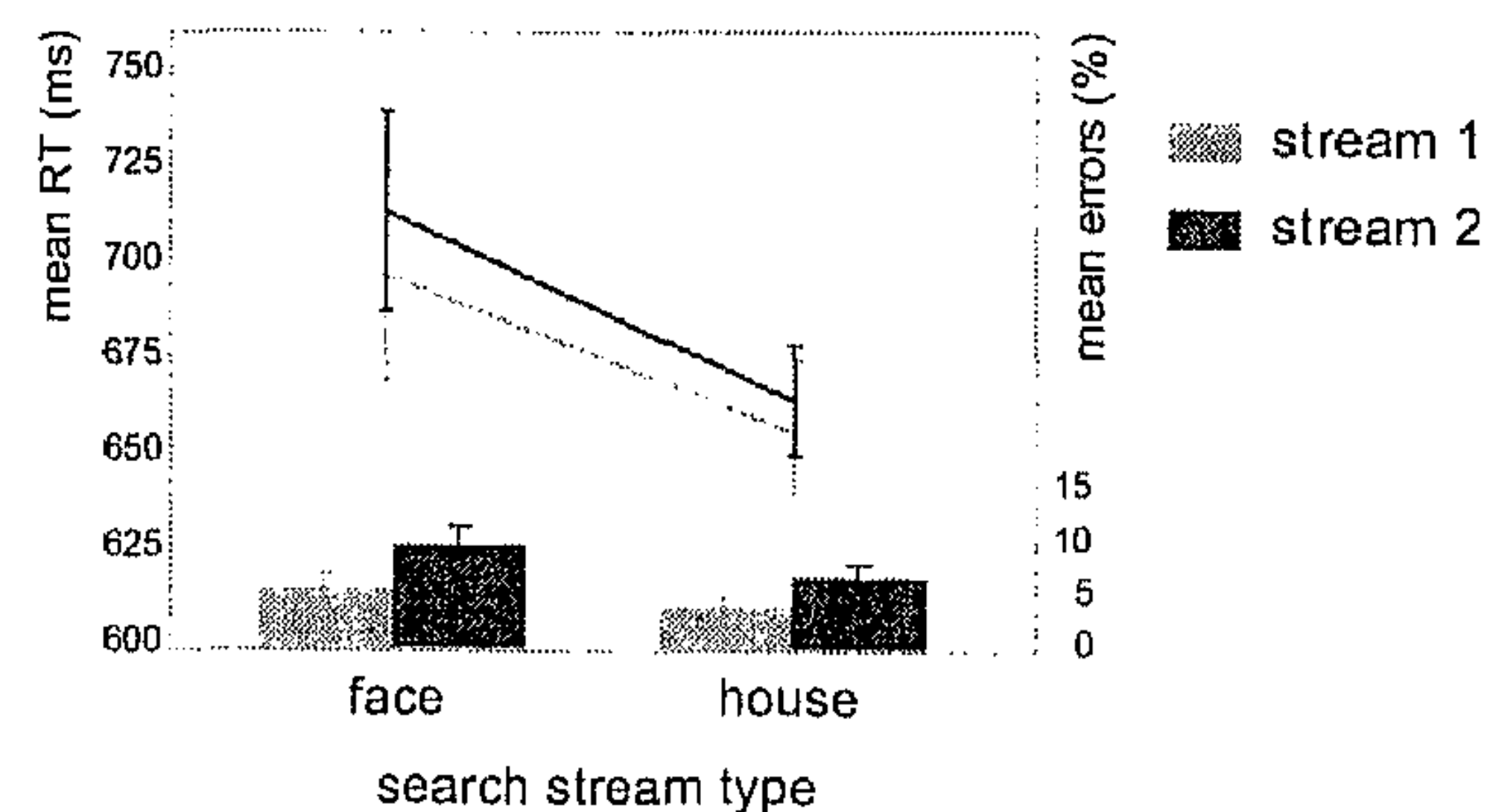
The FFA and PPA were identified for each participant in both hemispheres using an independent localizer run. To quantify potentially diverging responses to search targets, (regular) distractors, and memory items in these Volumes of Interest (VOIs), mean beta weights corresponding to the peak of the estimated hemodynamic responses were submitted to a two-way repeated measures ANOVA with hemisphere (left, right) and IOI type (search target, distractor, memory item of the same category, and memory item of a different category) as conditions.

Results

Behavioral results

On average, participants responded 684 ms (standard error ± 15 ms) after onset of a target stimulus. Participants failed to respond to a target (miss) or erroneously responded to the memory item, former target or (regular) distractor (false alarm) in 6.47% (standard error $\pm 1.23\%$) of the IOI presentations. Misses occurred more often than false alarms ($t(6) = 5.28$; $p < .003$). Memory items and former targets did not elicit more false alarms than other distractors (both streams; $ps > .9$). Although mean reaction time (RT) in the first and second search stream was similar ($F(1,6) = 3.12$; $p > .1$), participants made more errors in the second (8.26%) compared to the first (4.69%) search stream ($F(1,6) = 11.30$; $p < .02$; Figure 3). The performance on the second search stream was still good, however, indicating that the memory item was correctly remembered during the first search task.

Figure 3. Reaction time (RT) and accuracy for the two search stream types per stream. Error bars indicate standard errors of the mean (SEM).



Responses to items in face and house search streams did not differ, although there was a weak tendency towards a lower accuracy for face (92.30%) compared to house (94.75%) search streams (RT: $F(1,6) = 2.66$; $p > .1$; accuracy: $F(1,6) = 4.20$; $p > .05$). The lack of difference between responses in the two types of streams suggests that the differential transparency between house and face images was an effective manipulation to equalize task difficulty. No significant interactions between stream number and category were obtained for RT or accuracy.

fMRI results

Sustained neural modulations related to target search in the search stream

As depicted in Figure 4, multiple regression group analyses contrasting the search streams in which face or house search targets had to be detected, revealed strong bilateral activations in areas specialized in face and house processing, specifically the right FFA (Talairach coordinates (TC; in mm): 40, -46, -12), the left FFA (TC: -39, -47, -15), and the right (TC: 25, -37, -7) and left (TC: -24, -37, -9) PPA ($p(\text{Bonf}) < 4.5 \cdot 10^{-21}$). In addition, higher responses for house compared to face searches were revealed in the right lingual gyrus (TC: 4, -74, -6) and left middle occipital gyrus (TC: -28, -82, 25), whereas an area in the right middle temporal gyrus (TC: 45, -59, 7) showed the opposite response pattern. The event-related averages of FFA and PPA responses clearly reflect the strong and specific attentional modulation associated with, respectively face and house searches during the first as well as the second search stream (Figure 4). These differences between the two types of search streams are top-down driven, since sensory stimulation was identical in face and house search streams.

In contrast to the strong influence of the object category of the search target, no such top-down influence was revealed for the memory item. That is, even at more lenient thresholds, no activation differences were obtained by contrasting the first search streams during which a face or a house memory item was maintained in WM. Similar results were obtained when the different categories of the memory item were separately analyzed for the house and face search stream.

Widely distributed activity associated with identification of the target in the search stream

To reveal which brain areas were sensitive to target identification, neural responses to (correctly identified) search targets were contrasted with regular distractors. Target identification enhanced activation in a widely distributed network ($p(\text{Bonf}) < 0.05$; see Figure 5 and Table 1). Bilateral activations were found in visual areas, as well as in regions of the frontal and superior parietal cortex. Additional activations were revealed in the lateral sulcus, the cerebellum and in sub-cortical structures. Finally, left-lateralized enhancements were mainly revealed in the (post) central sulcus and cingulate cortex. These left-lateralized (presumably) button-press related activations were also consistently observed on the single-subject level, which supported the validity of the used rapid event-related fMRI design.

Conversely, no increased activations were revealed throughout the entire brain when memory items were contrasted to distractors (irrespective of whether memory items of the same or a different category were separately or jointly contrasted with distractors) even at more lenient thresholds.

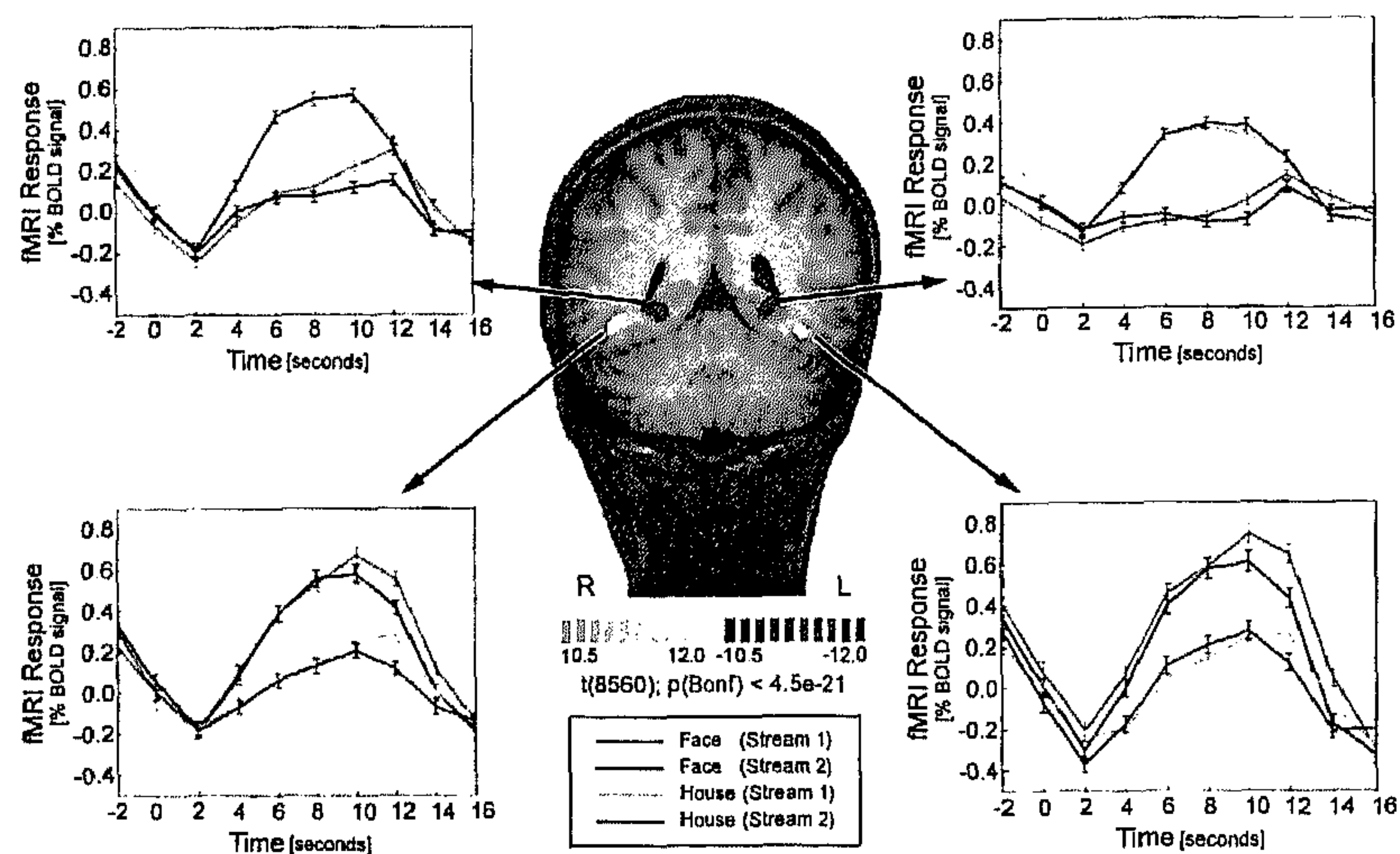


Figure 4. Results of contrasting the face search streams and house search streams at the group level shown on a participant's brain ($p(\text{Bonf}) < 4.5 \times 10^{-21}$; Bonf: Bonferroni corrected). Bilateral activity in the fusiform and parahippocampal gyrus is modulated by the type of category that is attended, as reflected in the event-related responses (averaged across subjects) during search streams in these clusters. Error bars indicate SEM. See text for details.

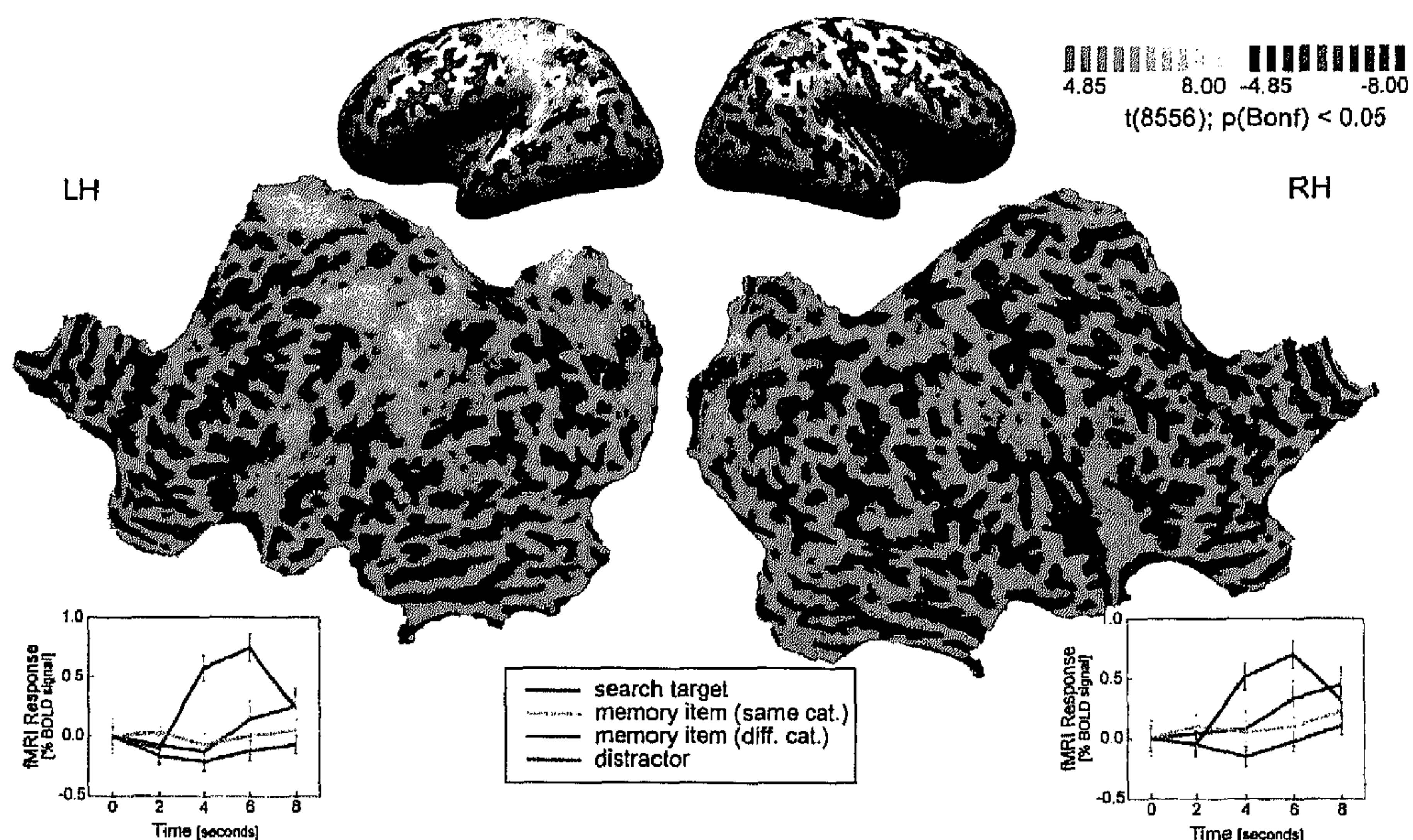


Figure 5. Cortical network involved in target identification projected on the flattened representation of a participant's grey-white matter boundary surface (LH: left hemisphere; RH: right hemisphere; dark grey: concave, light grey: convex cortical surface folding). Colored areas show a stronger activation in response to search targets compared to distractors in the first search stream, as identified by contrasting house and face targets with regular house and face distractors ($p(\text{Bonf}) < .005$; only activation patches with a size of more than 30 mm^2 are shown). Two examples of the estimated hemodynamic response functions (baseline corrected deconvolution analyses; see Figure 6 for details) in the left (left plot) and right (right plot) inferior frontal activation cluster (covering the inferior frontal gyrus, lateral sulcus, and insula; see Table 1 for details). Error bars indicate SEM.

area	left hemisphere				right hemisphere			
	x	y	z	size	x	y	z	size
LG	-9	-71	-4	3050	8	-56	-6	3901
FG	-	-	-	-	30	-47	-19	2455
Cu	-7	-74	10	6118	4	-77	10	5826
SPL/PrCu	-6	-59	32	17185	7	-60	31	10627
aIPL	-46	-55	32	2394	-	-	-	-
PCS/SMG/aIPS	-43	-35	47	11786	43	-41	44	4365
MTG	-56	-48	-3	1197	61	-38	1	774
pSTG/LATS	-57	-20	13	1170	-	-	-	-
TPJ/LATS	-54	-33	26	6261	55	-33	34	1503
PCS/LATS/STG	-43	-15	13	5509	52	-23	21	2149
IFG/LATS/INS	-40	1	10	5820	41	3	14	4524
CINGS	-4	-6	42	9110	4	-3	38	3553
aCINGS	-2	25	24	982	-	-	-	-
CS	-33	-21	54	6649	-	-	-	-
SFG/MFG	-32	32	36	2473	22	9	58	914
MFG	-27	54	15	1007	40	34	26	725

Table 1. Talairach coordinates (in mm) of the center of gravity and the size (in mm³) of clusters (with size > 700 mm³) activated in the group analysis, identified by contrasting house and face search targets with house and face distractors (see legend Figure 5 for details). Additionally, subcortical and cerebellar clusters were activated (not listed). Abbreviations: LG: lingual gyrus; FG: fusiform gyrus; Cu: cuneus; SPL: superior parietal lobule; PrCu: precuneus; IPL: inferior parietal lobule; PCS: postcentral sulcus; SMG: supramarginal gyrus; IPS: intraparietal sulcus; STG: superior temporal gyrus; STG: middle temporal gyrus; LATS: lateral sulcus; TPJ: temporo-parietal junction; IFG: inferior frontal gyrus; INS: insula; CINGS: cingulate sulcus; CS: central sulcus; SFG: superior frontal gyrus; MFG: middle frontal gyrus; a: anterior; p: posterior.

Differential responses to search items revealed by deconvolution analyses in FFA and PPA

To estimate responses to individual IOIs presented in the face and house search streams, deconvolution analyses in subject-specific VOIs were performed. For each subject, the FFA and PPA were determined in both hemispheres (Table 2) based on an independent localizer run. The estimated hemodynamic responses to the four IOI types in FFA and PPA are shown for the first stream in Figure 6. The estimated hemodynamic response peaked 4-6 s after stimulus onset, which is consistent with peak latencies observed in other event-related studies. Peak amplitudes (Figure 7) differed between the various IOI types in the FFA ($F(1,6) = 4.15$; $p < .03$) and tended to differ in the PPA ($F(1,6) = 2.52$; $p < .10$). Planned comparisons showed that in both VOIs, responses to search targets were stronger than responses to regular distractors (FFA: $t(6) = 2.73$; $p < .04$; PPA: $t(6) = 6.24$; $p < .001$).

In contrast, the memory items did not evoke different responses compared to regular distractors, irrespective of whether the memory items belonged to the same (FFA: $t(6) = 0.42$; $p > .65$; PPA: $t(6) = 1.56$; $p > .15$) or to a different (FFA: $t(6) = 0.03$; $p > .95$; PPA: $t(6) = 0.95$; $p > .35$) category as the search target¹. No main effects of hemisphere (FFA: $F(1,6) = 0.52$; $p > .45$; PPA: $F(1,6) = 0.35$; $p > .50$) or interactions between IOI type and hemisphere (FFA: $F(1,6) = 0.47$; $p > .70$; PPA: $F(1,6) = 0.01$; $p > .90$) were revealed.

¹ Figure 7b suggests that the difference between the memory item and distractor condition might become significant, if standard errors would be smaller. However, inspection of individual data revealed that only a few subjects contributed to this potential effect. For example, the mean peak amplitudes of the house memory item (mean = 0.44; sem: 0.1) and the distractor (mean = 0.41; sem: 0.1) were almost similar when two subjects were excluded from the analyses. Likewise, the mean peak amplitudes of the face memory item (mean = 0.40; sem: 0.2) and the distractor (mean = 0.38; sem: 0.1) were almost identical when one subject was left out.

Table 2. Talairach coordinates (in mm) of the center of gravity and the size (in mm³) of the VOIs identified for each subject.

#subject	FFA								PPA							
	left hemisphere				right hemisphere				left hemisphere				right hemisphere			
	x	y	z	size	x	y	z	size	x	y	z	size	x	y	z	size
1	-40	-42	-19	500	37	-50	-22	1054	-26	-49	-9	3576	26	-40	-12	3370
2	-39	-50	-10	286	41	-43	-8	223	-27	-56	-7	396	20	-41	-7	1355
3	-43	-54	-20	52	39	-49	-18	285	-24	-39	-9	247	25	-34	-11	501
4	-38	-48	-14	207	36	-59	-15	405	-24	-39	-6	860	22	-40	-4	1191
5	-34	-58	-11	274	37	-44	-11	286	-26	-39	-7	2303	25	-37	-7	2657
6	-39	-45	-14	104	32	-59	-12	222	-25	-50	-4	411	23	-46	-7	1663
7	-33	-44	-14	1158	39	-43	-13	1256	-22	-49	-5	2311	26	-49	-5	2802
mean	-38	-49	-15	369	37	-50	-14	533	-25	-46	-7	1443	24	-41	-7	1934

Neural activity related to updating the search template

During presentation of the second cue (i.e., the cue in between the two search streams) participants had to update the search template, since the target of the second search stream (i.e., the task-irrelevant memory item in the first search stream) would become relevant whereas the target of the first search stream should further be ignored. As shown in Figure 8, presentation of a cue indicating that the target of the second search stream would become relevant activated a widely distributed network of brain areas including, among others, regions in the middle frontal gyrus (MFG), the anterior insula, the pre-supplementary motor area (pre-SMA), and the posterior parietal cortex (PPC; see Table 3 for details).

area	left hemisphere				right hemisphere			
	x	y	z	size	x	y	z	size
aINS	-31	18	8	2357	40	15	7	5754
MFG	-41	21	31	10125	39	22	36	15254
pre-SMA	-5	6	46	4374	8	17	38	2964
PrCS	-39	-7	51	5873	-	-	-	-
TPJ/IPL	-52	-42	27	2312	52	-35	19	3225
STG	-52	-46	12	8052	54	-40	8	4882
IPL	-38	-47	43	5421	40	-46	42	6607
IPS/SPL	-28	-62	46	4344	28	-62	41	4019

Table 3. Talairach coordinates (in mm) of the center of gravity and the size (in mm³) of clusters (with voxel size > 200 mm³) obtained by the conjunction analysis of the second house cue and second face cue compared to rest (see legend Figure 8 for details). Additional to these clusters, clusters in the precuneus, cuneus, and other parts of the visual cortex were activated, as well as subcortical clusters. Abbreviations: aINS: anterior insula; MFG: middle frontal gyrus; pre-SMA: pre-supplementary motor area; PrCS: precentral sulcus; TPJ: temporoparietal junction; IPL: inferior parietal lobule; STG: superior temporal gyrus; IPS: intraparietal sulcus; SPL: superior parietal lobule.

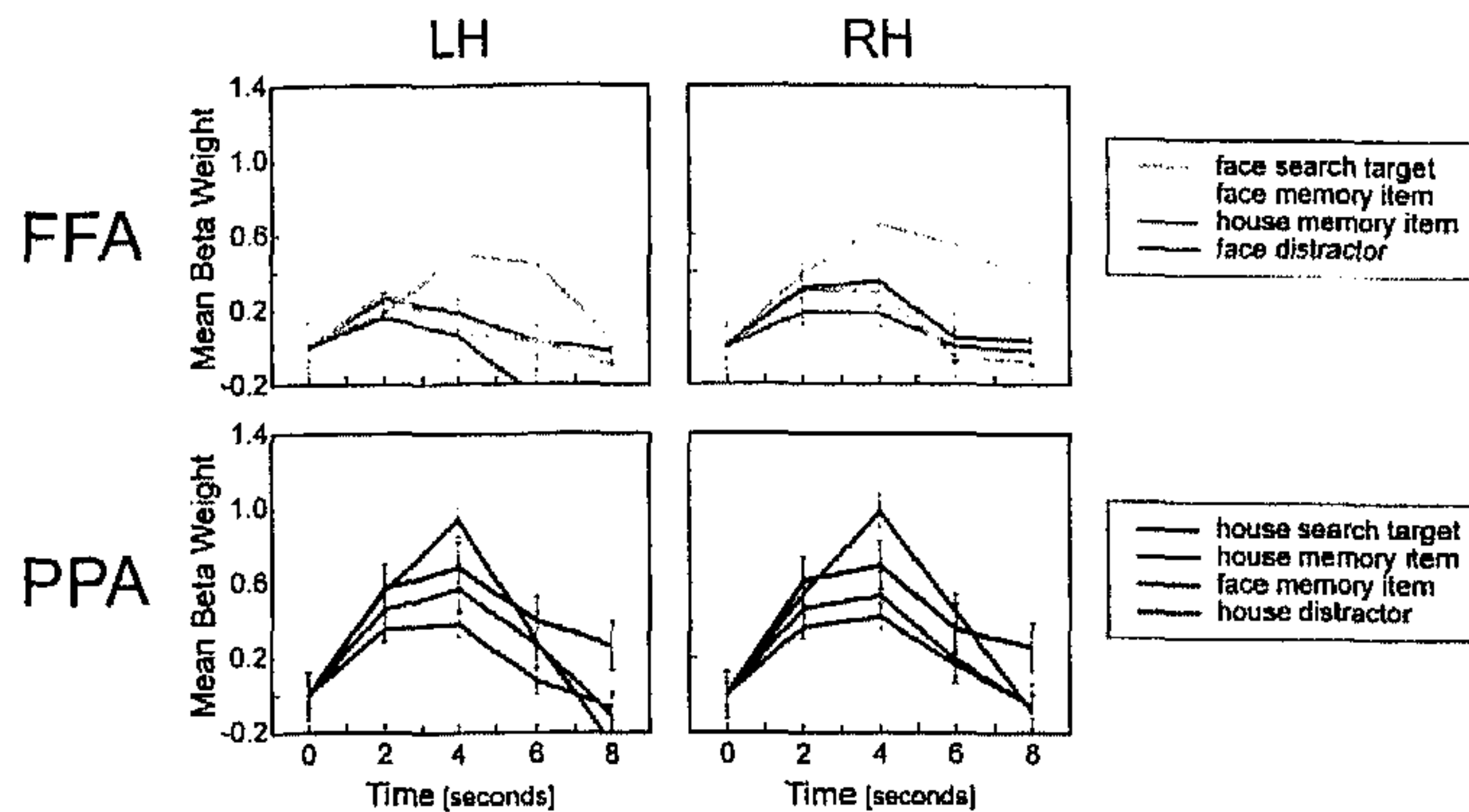


Figure 6. Hemodynamic response functions for the four IOI types presented in the first face (for FFA; upper row) or first house (for PPA; lower row) search stream estimated for each data point (i.e., each volume acquisition = 2 s) independently by the deconvolution analysis. The onset of the IOI presentation was at 0 s (data point 1). To correct for potential differences in baseline, this first data point was subtracted from all data points of each individual function. Error bars indicate SEM.

Figure 7. Mean peak amplitudes of the hemodynamic response functions in a). FFA and b). PPA. Amplitudes correspond to the mean response at 4-6 s (for FFA) or 4 s (for PPA) averaged over hemispheres. See Figure 6 for details. Error bars indicate SEM.

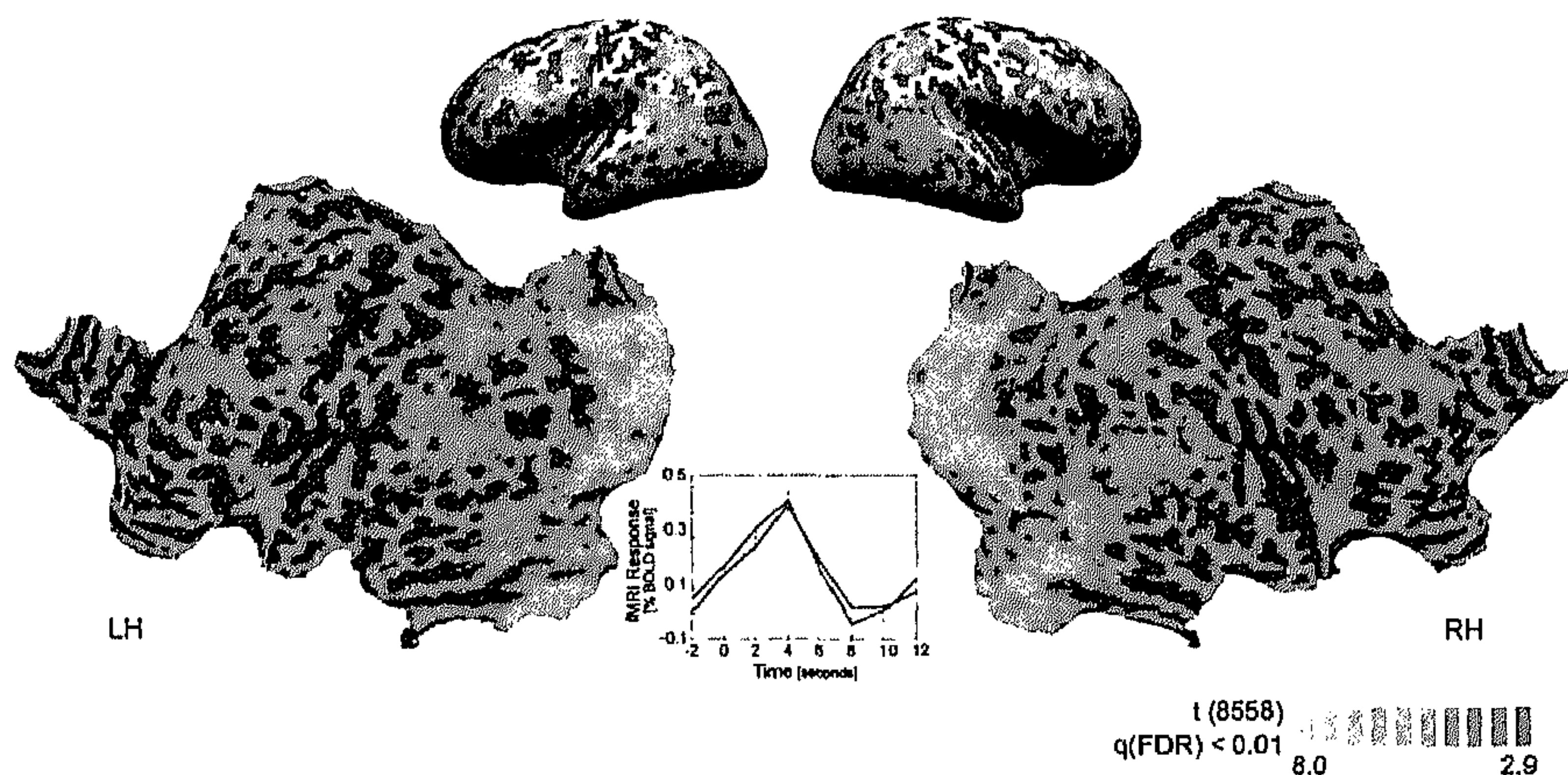
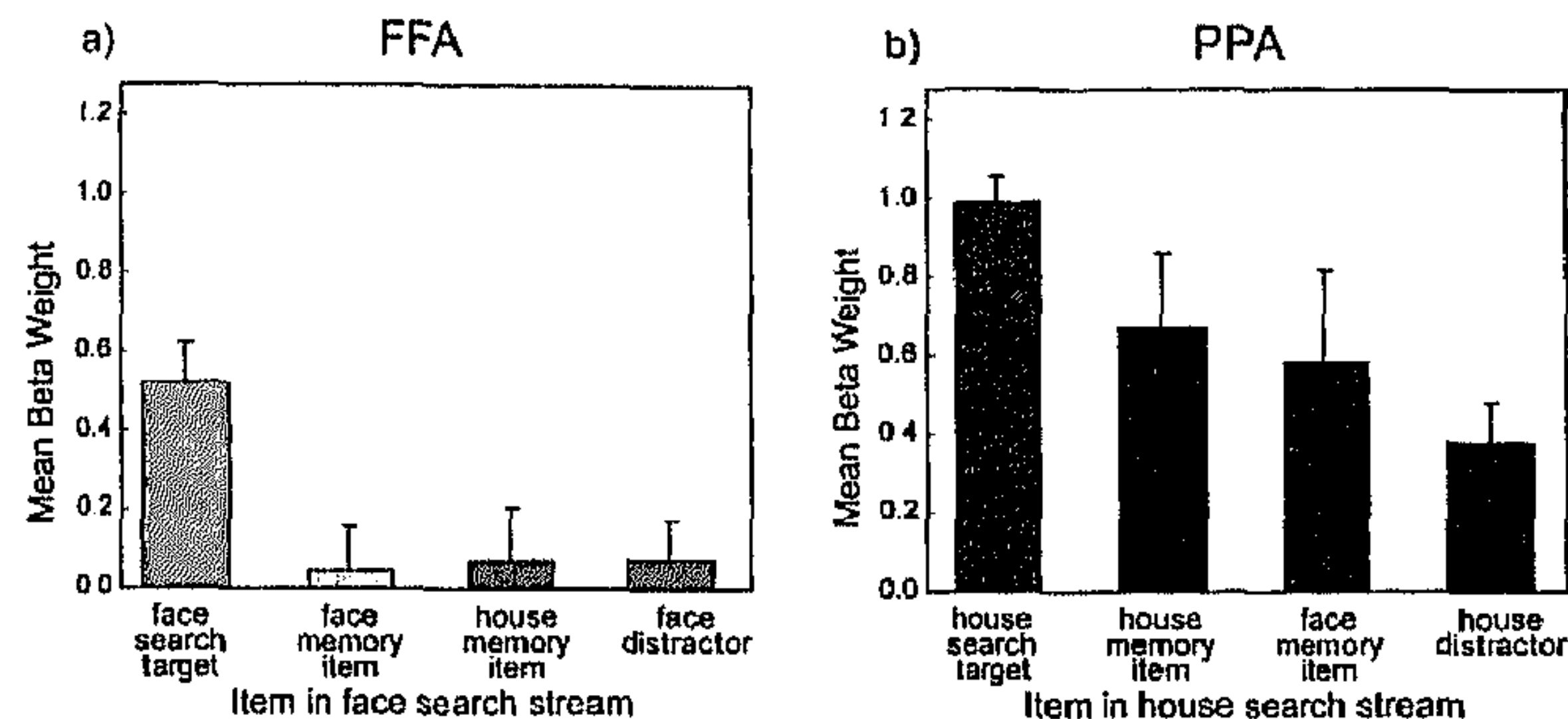


Figure 8. Network involved in updating search template information in WM projected on the flattened representation of a participant's grey-white matter boundary surface (LH: left hemisphere; RH: right hemisphere; dark grey: concave, light grey: convex cortical surface folding). Colored areas show a stronger activation during presentation of the second cue (i.e., "1" or "2"; see Figure 2) compared to rest (i.e., intertrial interval) as resulting from a conjunction analysis in which the cues for second stream house and face targets were separately contrasted with rest (False Discovery Rate: $q(\text{FDR}) < .001$; only positive activations with a size of more than 200 mm² are shown). An example of an event-related average for the activated region in the middle frontal gyrus of the right hemisphere, showing transient prefrontal activity related to updating the search template (cue indicating an upcoming face [blue] or house [red] search stream), is shown in the insert. Error bars indicate SEM.

Discussion

The present event-related fMRI study aimed to assess the influence of WM content on attentional search processes, by requiring participants to search a face or house target in a stream of superimposed face and house images, while maintaining another face or house in WM which functioned as target in a subsequently presented second stream. By examining sustained neural responses during search, as well as by investigating transient responses to memory items presented in the first stream, attentional guidance by WM content other than the search template could be investigated.

Behavioral results showed that although participants made more errors in the second compared to the first stream, performance on the second search stream was still good. This indicates that the memory item was correctly remembered during the first search task. Moreover, the false alarm rate caused by the memory items was similar to the rate caused by regular distractors that were not represented in memory. Thus, our subjects were well able to memorize two objects while using only one of them as a template during visual search. Finally, performance on house and face search streams was equally good, suggesting that the search tasks had a similar difficulty.

The search template - but not the accessory memory item - induces category-specific sustained attentional enhancements in FFA and PPA

Category-specific sustained attentional modulations were revealed in the FFA and PPA throughout visual search in the first as well as the second search stream. The higher FFA and PPA responses during respectively face and house search are consistent with previous fMRI studies (e.g., O'Craven, Downing, & Kanwisher, 1999; Serences et al., 2004) showing that attending objects of a certain category results in enhanced activity of higher visual areas specialized in processing objects of that category. The category-specific modulations in the present study could not be stimulus-driven, since stimulus input did not differ between both search types. Rather, top-down influences, signaling information concerning the object category of the search target, appear to drive the enhanced processing of the search item's attribute corresponding to this category. That is, when the search target was a face the face attribute of search items was more intensively processed, whereas attention was biased towards the house attribute when the search target was a house. This finding is in accordance with a study that simultaneously recorded activity in the prefrontal (PFC) and inferior temporal cortex of monkeys performing a delayed match-to-category task (Freedman, Riesenhuber, Poggio, & Miller, 2003). Results showed that the PFC cortex maintained task-relevant information on a category-specific level (see also Freedman, Riesenhuber, Poggio, & Miller, 2001), which appeared to be fed back to inferior temporal cortex when the category of the sample stimulus had to be matched to that of the test stimulus.

The main focus of the present study was, however, on the feedback of information concerning the memory item instead of the search target. No activation differences were obtained throughout the brain when a face compared to a house memory item was maintained in WM during visual search, suggesting that only the search target and not the memory item influenced to which attribute of the search item attention was deployed.

Transient match enhanced responses to search targets - but not to memory items

Correct identification of target items in the search stream was associated with enhanced activation in a widely distributed network. This enhanced processing of stimuli matching the search template is consistent with the “match enhancement” observed in monkey prefrontal and inferior temporal cortex (Miller & Desimone, 1994; Miller, Erickson, & Desimone, 1996). In addition, the revealed enhancements are in line with target-related activations obtained in a previous fMRI study using a similar temporal search task (Jiang et al., 2000). However, although the present network partly overlapped with the network obtained by Jiang et al. (2000), it was more extended including additional areas in mainly the parietal and prefrontal cortex. This additional recruitment of areas probably resulted from the increased difficulty to detect targets in the present task compared to the task of Jiang et al. (2000), in which face (instead of superimposed house-face) images were presented at a much slower rate.

Importantly, no match enhancements were observed in our study when memory items were encountered in the search stream. However, the overlap between neural responses in our rapid event-related design might have obscured smaller transient increases elicited by the memory item. Therefore, deconvolution analyses were carried out to minimize the interference between responses to temporally adjacent events (Glover, 1999; Serences, 2004). Deconvolution analyses were confined to the FFA and PPA, which were independently mapped for each subject. Consistent with the multiple-regression results, FFA and PPA responses were enhanced for processing of search targets compared to distractors. Conversely, no indications for “match enhanced” responses were obtained for the memory item, irrespective of whether the memory item belonged to the same or to a different object category as the search target. This suggests that memory items do not receive more attention than distractors not represented in memory, consistent with the results of our ERP study (Chapter 2).

Results of neurophysiological studies in non-human primates have shown that prefrontal activity reflects target-specific information (Rainer, Asaad, & Miller, 1998), which can vary according to the dimension of the stimulus that must be attended (Sakagami & Niki, 1994) or the task that needs to be performed (Asaad, Rainer, & Miller, 2000), indicating that PFC neurons selectively encode and maintain task-relevant information. In the present study both the search target as well as the memory item should be encoded and maintained in WM, since they

are either currently relevant or will become relevant in the near future. Nevertheless, although both items are held in memory, results of the multiple-regression as well as the deconvolution analyses indicated that search items were only matched to the search template but not to the accessory memory item. How is this top-down control regulated, such that only the search target representation - and not the other accessory memory items - serves as search template and provides feedback to visual areas? Moreover, how can the roles of these WM representations be switched in the second part of the task, in which the memory item functions as search template? Several cell recording studies have suggested that the PFC represents the behavioral context in a highly flexible manner, rather than that it passively maintains sensory information (for review, see Miller, 2000). For example in a delayed paired associate task, PFC delay activity shortly reflects the sample immediately after its presentation, but subsequently conveys prospective information concerning the anticipated object associated with the sample (Rainer, Rao, & Miller, 1999). Moreover, another study showed that when the identity of an object first has to be remembered followed by its location, many PFC neurons “switch” from conveying identity information in the first part to reflecting location information in the second part of the task, in accordance with changing task demands (Rao, Rainer, & Miller, 1997). We suggest that in a similar flexible manner, the PFC dynamically controls at which point in time a certain representation in WM is allowed to provide feedback to visual areas during search.

Updating of the search template in working memory

Although the present study mainly focused on neural responses during visual search, additional analyses were performed to study activations in between the first and second search stream when the search template in WM needed to be updated. Updating search template information engaged a distributed network of areas including, among others, regions in the MFG, pre-SMA, and PPC. This network closely resembled the neural activation patterns observed in a recent fMRI study (Roth, Serences, & Courtney, 2006), in which participants had to update information in WM concerning a house (or face) target that had to be searched in a RSVP of houses (or faces). Similar activations were, for example, obtained in the middle part of the MFG (Brodmann area 46). The recruitment of this area during the updating of the search template in WM is consistent with its role in manipulating WM content, such as when items in WM need to be reordered (e.g., D' Esposito, Postle, Ballard, & Lease, 1999) or selected (Rowe & Passingham, 2001). Furthermore, the updating network in the present study showed overlap with networks that are recruited when the relevancy between two representations in WM is reversed (Garavan, Ross, Li, & Stein, 2000; Sylvester et al., 2003), a feature which is also present in our task.

Only the search template guides attentional selection

As expected, several neural indications of top-down attentional guidance by the search template have been revealed in the present study: First, searching a specific target object in a stream of stimuli was associated with sustained attentional enhancements in visual areas specialized in processing the object category to which the search target belonged. This indicated that category-specific target information is signaled to these visual areas promoting the enhanced processing of search item attributes corresponding to this category. Second, identification of the target increased activity in a widely distributed network, probably reflecting the detection of a match between top-down feedback and bottom-up input. This match enhancement was also clearly present in higher visual areas, suggesting that visual input matching the search template is attentively processed. These findings are consistent with previous fMRI results showing selective processing of the house or face attribute of compound stimuli (Serences et al., 2004) and enhanced processing of target stimuli (Jiang et al., 2000). Moreover, the top-down feedback from WM concerning category- as well as item-specific target information is in line with neurophysiological findings suggesting that the PFC is able to convey both category- (Freedman et al., 2001, 2003) as well as item-specific (Rainer, Asaad, & Miller, 1998) information.

The primary focus of the study was, however, on the attentional guiding capabilities of accessory memory items. Examination of sustained and transient neural modulations during visual search did not reveal any indications for attentional guidance by the accessory memory item. This suggests that the top-down feedback signal to which visual input is matched only carries information concerning the search template but does not convey information concerning the memory item. However, when the first search stream is finished, top-down feedback should provide information concerning the memory item (which has become the new search template) and not about the target of the first stream (which does not function as search template anymore). Updating the search template from the first stream to the second stream target engaged a (predominantly) fronto-parietal network. Presumably, parts of this network are involved in maintaining the target representations as well as in controlling which representation is allowed to provide feedback to lower areas during visual search. The MFG, which shows match enhanced responses to targets presented in the search stream (Figure 5) and which is also recruited when the search template needs to be updated (Figure 8), might play an important role in controlling that only the search template provides feedback during visual search. This suggestion is supported by findings that lesions in the MFG result in stronger attentional guidance by memory items (Soto, Humphreys, & Heinke, 2006).

Although additional research is needed to establish which areas are involved and how they are able to regulate top-down control by WM representations, one might speculate on several potential mechanisms that could subserve this function (following previous suggestions by Downing & Dodds,

2004). To avoid interference between items in WM, the search template and the accessory memory items might be stored in separate partitions of the PFC, similar to the separate storage of previous and future goals in monkey PFC (Genovesio, Brasted, & Wise, 2006). Alternatively, they might be stored in the same partition, in which they are in competition for providing feedback to visual areas. In the latter case, neurons reflecting ‘task context’ (e.g., Wallis, Anderson, & Miller, 2001; White & Wise, 1999; see also Miller & Cohen 2001) might bias competition between WM representations in favor of the task-relevant search template, whereas irrelevant representations are suppressed.

Conclusion

Results of the present event-related fMRI study showed that top-down feedback from the search template in WM enhanced processing of stimuli belonging to the same object category and of matching items. In contrast, the accessory memory item did not induce such category- or item-specific enhancements, indicating that only the search template guided attention during visual search. Only when the memory item became relevant in the second half of the task, it was able to guide attention. This suggests that attentional guidance by WM content is highly selective, as exclusively task-relevant feedback is provided to visual areas.

Acknowledgments

This research was supported by NWO Grant No. 402-01-632 to JCP and RG. Joel Reithler and Bettina Sorger are acknowledged for their contributions during several stages of this project.

References

- Asaad, W. F., Rainer, G., & Miller, E. K. (2000). Task-specific neural activity in the primate prefrontal cortex. *Journal of Neurophysiology*, 84, 451–459.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, 97, 523–547.
- Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, 80, 2918–2940.
- Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, 363, 345–347.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.
- D'Esposito, M., Postle, B. R., Ballard, D., & Lease, J. (1999). Maintenance versus manipulation of information held in working memory: an event-related fMRI study. *Brain & Cognition*, 41(1), 66–86.

- Downing, P. E., & Dodds, C. M. (2004). Competition in visual working memory for control of search. *Visual Cognition*, 11(6), 689-703.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96(3), 433-458.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676), 598-601.
- Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, 291(5502), 312-316.
- Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2003). A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience*, 23(12), 5235-5246.
- Fuster, J. M., Bauer, R. H., & Jervey, J. P. (1985). Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Research*, 330(2), 299-307.
- Garavan, H., Ross, T. J., Li, S. J., & Stein, E. A. (2000). A parametric manipulation of central executive functioning. *Cerebral Cortex*, 10(6), 585-592.
- Genovesio, A., Brasted, P. J., & Wise, S. P. (2006). Representation of future and previous spatial goals by separate neural populations in prefrontal cortex. *Journal of Neuroscience*, 26(27), 7305-7316.
- Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage*, 9(4), 416-429.
- Houtkamp, R., & Roelfsema, P. R. (2006). The effect of items in working memory in the deployment of attention and the eyes during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 32(2), 423-442.
- Jiang, Y., Haxby, J. V., Martin, A., Ungerleider, L. G., & Parasuraman, R. (2000). Complementary neural mechanisms for tracking items in human working memory. *Science*, 287(5453), 643-646.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17, 4302-4311.
- Miller, E. K. (2000). The neural basis of top-down control of visual attention in prefrontal cortex. In: S. Monsell, & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance* (vol. XVIII; pp 511-534). Cambridge, MA: MIT Press.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal function. *Annual Review of Neuroscience*, 24, 167-202.
- Miller, E. K., & Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science*, 263(5146), 520-522.
- Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, 16(16), 5154-5167.
- O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as units of attentional selection. *Nature*, 401, 584-587.
- Oh, S. H., & Kim, M. S. (2003). The guidance effect of working memory load on visual search. *Journal of Vision*, 3(9), 629a.
- Olivers, C. N., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 32(5), 1243-1265.
- Rainer, G., Asaad, W. F., & Miller, E. K. (1998). Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature*, 393, 577-579.
- Rainer, G., Rao, S. C., & Miller, E. K. (1999). Prospective coding for objects in the primate prefrontal cortex. *Journal of Neuroscience*, 19, 5493-5505.
- Rao, S. C., Rainer, G., & Miller, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, 276, 821-824.
- Roth, J. K., Scerences, J. T., & Courtney, S. M. (2006). Neural system for controlling the contents of object working memory in humans. *Cerebral Cortex*, 16, 1595-1603.

Chapter 3

- Sakagami, M., & Niki, H. (1994). Encoding of behavioral significance of visual stimuli by primate prefrontal neurons: Relation to relevant task conditions. *Experimental Brain Research*, 97, 423–436.
- Rowe, J. B., & Passingham, R. E. (2001). Working memory for location and time: activity in prefrontal area 46 relates to selection rather than maintenance in memory. *Neuroimage*, 14, 77–86.
- Sanders, A. F. (1990). Issues and trends in the debate on discrete vs. continuous processing of information. *Acta Psychologica*, 74, 123–167.
- Serences, J. T. (2004). A comparison of methods for characterizing the event-related BOLD timeseries in rapid fMRI. *Neuroimage*, 21(4), 1690–1700.
- Serences, J. T., Schwarzbach, J., Golay, X., Courtney, S. M., & Yantis, S. (2004). Control of object-based attention in human cortex. *Cerebral Cortex*, 14, 1346–1357.
- Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 31(2), 248–261.
- Soto, D., Humphreys, G. W., & Heinke, D. (2006). Dividing the mind: The necessary role of the frontal lobes in separating memory from search. *Neuropsychologia*, 44, 1282–1289.
- Sylvester, C. Y., Wager, T. D., Lacey, S. C., Hernandez, L., Nichols, T. E., Smith, E. E., & Jonides, J. (2003). Switching attention and resolving interference: fMRI measures of executive functions. *Neuropsychologia*, 41(3), 357–370.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. Stuttgart: G. Thieme.
- Tomita, H., Ohbayashi, M., Nakahara, K., Hasegawa, I., & Miyashita, Y. (1999). Topdown signal from prefrontal cortex in executive control of memory retrieval. *Nature*, 401, 699–703.
- Wallis, J. D., Anderson, K. C., & Miller, E. K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, 411(6840), 953–956.
- White, I. M., & Wise, S. P. (1999). Rule-dependent neuronal activity in the prefrontal cortex. *Experimental Brain Research*, 126(3), 315–335.
- Woodman, G. F., & Luck, S. J. (in press). Do the contents of visual working memory automatically influence attention selection during visual search? *Journal of Experimental Psychology: Human Perception and Performance*.

Chapter 4

Scaling the “zoom-lens” of attention

on: Peters, J. C. & Goebel, R. Scaling the “zoom-lens” of attention. (*in*
ration)

Abstract

Although it has been shown that humans can modify the size of their spatial attention window, it remains unclear how the process of adjusting this attentional “zoom-lens” is implemented. We performed a series of four experiments to explore this scaling mechanism using a newly developed three-level compound stimulus. Response congruency between the different levels was varied to assess the degree of interference between levels, when attention was scaled to a single level and, in some trials, subsequently rescaled to another level within the same stimulus. The first study suggested that attention could be accurately scaled to each level separately, which indicated that the attentional distribution could have an annular shape. In addition, the second experiment showed that the middle level interfered when attention was rescaled between the large and the small level, indicating that the zoom-lens could not be discretely rescaled between the two levels. The third experiment revealed that both the “direction” (i.e., zooming-in versus zooming-out) as well as the extent of rescaling influenced the way in which the attentional zoom-lens was adjusted. The results of this experiment additionally suggested that under optimal circumstances, discrete scaling of the zoom-lens does appear to be possible. Finally, a fourth experiment showed that the partial occlusion in the used compound stimulus did not affect the obtained results. Overall, the present findings indicate that attentional scaling mechanisms are dynamically influenced by both top-down factors, like strategy, as well as bottom-up factors, such as visual context.

Introduction

There is ample evidence that visual attention can be oriented to a specific location in the visual field, thereby facilitating the processing of stimuli presented at that particular location (e.g., Eriksen & Hoffman, 1974; Posner, 1980). Besides this ability to shift the focus of attention through the visual field, the zoom-lens theory of attention (Eriksen & St. James, 1986) suggests that the *size* of the focus can be adjusted to match the spatial scale of a relevant region. That is, the size of the attentional window can be constricted or expanded to focus attention, for example, on a single letter or a whole word within a given sentence (LaBerge, 1983). According to the zoom-lens theory, a tradeoff exists between the size of the zoom-lens and attentional resolution (Eriksen & St. James, 1986), due to limited attentional capacity. Several studies have suggested that the zoom-lens could indeed be constricted or expanded according to task demands, whereas the size of the zoom-lens was negatively correlated with processing efficiency (Castiello & Umiltà, 1990; Eriksen & St. James, 1986; LaBerge, 1983; Luo, Greenwood, & Parasuraman, 2001; McMains & Somers, 2005; Müller, Bartelt, Donner, Villringer, & Brandt, 2003a). In addition to shifting the attentional focus, scaling the focus also appears to play a vital role in visual search (Greenwood & Parasuraman, 1999). However, compared to the vast amount of studies investigating covert attention shifts, relatively few experiments have been conducted on the scaling of covert attention. For example, several studies explored whether the attentional focus is shifted from one location to another in an analog (i.e., covering other locations along the way) or discrete (i.e., ‘jumping’ from the first to the second location) manner, and likewise, whether the time required to perform an attention shift depends on the distance over which attention has to be shifted (e.g., Chastain, 1992a, 1992b; Murphy & Eriksen, 1987; Shulman, Remington, & MacLean, 1979; Sperling & Weichselgartner, 1995; Tsal, 1983). However, to our knowledge these types of questions have not been studied in the context of the zoom-lens model. For example, it is unclear whether the zoom-lens is scaled in a continuous or discrete fashion. Likewise, whether there is a relation between the time needed to rescale and the extent of rescaling is unknown.

To study these and other issues concerning the scaling of the attentional zoom-lens, we set out to perform a series of experiments using a compound stimulus with three levels. Initially, the intention was to extend the classical Navon figures (Navon, 1977) by creating a three-level hierarchical compound stimulus in which the letters on the different levels would have a similar identification difficulty (to avoid confounding differences in figural quality between levels; Hofman, 1980). Since pilot results suggested that this was extremely difficult, we decided to use another type of three-level compound stimulus (Vorberg, 1992, unpublished; Figure 1) that shares some characteristics, but differs in others from the classical hierarchical compound stimulus: The primary difference is that one level of the compound stimulus is not constituted from elements of another level.

Although this might complicate comparisons with studies using Navon figures, the presence of only one element at the middle and small level facilitate a better control of attentional rescaling behavior, which is important for investigation of our hypotheses. Another difference is the partial occlusion of the middle and large level in the new compound stimulus. However, potential confounds by occlusion were controlled by using a dynamic version of the compound stimulus in one of the experiments (experiment 4) discussed below.

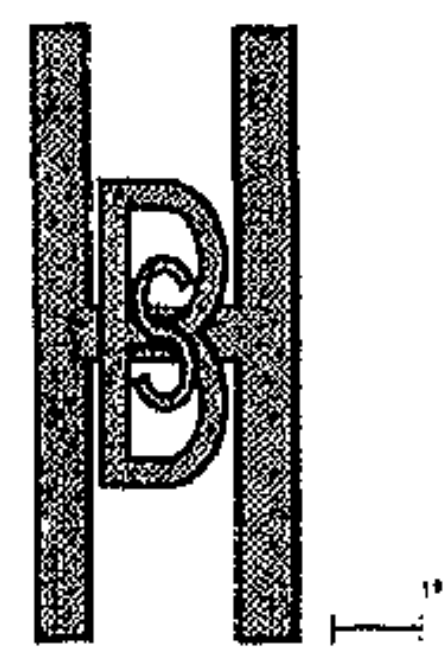


Figure 1. Example of the three-level compound stimulus used in the present experiments. One of the levels was cued in advance by a outlined rectangle, demarcating the position of the cued level. By scaling attention according to this cue the identity of the letter at the cued level could be reported as fast and accurately as possible. In experiment 2 and 3, a non-letter character was presented at the cued level in some of the trials, indicating that attention had to be rescaled to another level within the same stimulus. The stimulus is drawn to scale.

Several research questions concerning the scaling of the attentional zoom-lens could be addressed by using this type of compound stimulus: First, the accuracy with which the zoom-lens could be scaled to the different levels was explored in experiment 1, by cueing participants to one of the three levels to perform a letter identification task. Response congruency between the different levels was manipulated to study whether attention could be exclusively deployed to a single level. Based on findings of studies using the flanker paradigm (e.g., Eriksen & Eriksen, 1974), we hypothesized that if the zoom-lens can not be solely scaled to one level but also unintentionally covers other levels, response incongruency between these levels would lead to behavioral response congruency costs (RCCs). However, several studies have suggested that the zoom-lens can have an annular shape (Egley & Homa, 1984; Juola, Bouwhuis, Cooper, & Warner, 1991; Müller & Hübner, 2002; see Chapter 7 of this thesis). This implies that in the three-level stimulus, participants should be able to attend the middle or large level, without attending the small level, even though the middle of the stimulus was fixated. Secondly, the mechanisms underlying scaling of the zoom-lens were investigated in experiment 2 and 3, by requiring participants to rescale attention from the cued level to a different level. The primary goal here was to uncover whether the zoom-lens is continuously or discretely rescaled. Therefore, the experiment included trials in which attention had to be rescaled from the small to the large level and vice versa. If the zoom-lens were rescaled continuously between the large and small level, the zoom-lens should temporarily encompass the intermediate letter. This would lead to RCCs if the letters at the intermediate level and at the level to which attention had to be rescaled (further referred to as target level) were coupled to a different instead of the same response. In contrast, if the zoom-lens were rescaled discretely, attention would not be deployed to the intermediate level and response congruency effects should be (close to) absent. To ensure that solely covert attention played a role in rescaling, eye movements were monitored in experiment 2. In a third experiment, the divergent effect of expanding versus

contracting the zoom-lens (as observed in experiment 2) was studied in more detail. In addition, experiment 3 assessed whether the extent of attentional rescaling influenced performance. Finally, to ensure that the effects obtained in the present experiments were not influenced by the differences in occlusion between the different levels, a fourth experiment was performed using a ‘dynamic’ version of the compound stimulus in which none of the three levels was occluded by another level.

Experiment 1: how accurately can attention be scaled to each of the three levels?

The first experiment studied the accuracy with which the zoom-lens could be scaled to each of the levels in the three-level compound stimulus (Figure 1) without encompassing (parts of) the other levels. An ‘annular’² distribution of attention would be optimal to exclusively attending the large (or to the middle) level without interference of the other levels. As mentioned above, results of several studies (e.g., Juola et al., 1991; Müller, & Hübner, 2002) suggested that attention can indeed be deployed to a ring-shaped region, while stimuli in the center are ignored. However, in the present design the width of the annulus should be relatively small when attending the large level, to minimize response congruency effects of not only the small but also the middle level. To estimate the minimal width of the annulus, a compound stimulus version with a relative large (labeled largeMiddle or *laM* stimulus) and one with a relative small (*smM* stimulus) size of the middle letter was used. We hypothesized that attending the large level in the *laM* stimulus would lead to more interference than in the *smM* stimulus, since the distance between the large and middle level is smaller in the first compared to the second stimulus. Likewise, if the size of the zoom-lens could not be constricted to encompass solely the small level, attending the small level in the *laM* stimulus would lead to less interference than in the *smM* stimulus, since the distance between the middle and small level is larger in the *laM* compared to *smM* stimulus.

Methods

Participants

Twelve healthy volunteers (4 males; mean age: 21.4 years; standard error: 1.0 years) with normal or corrected-to-normal visual acuity were paid for participation in this study. All participants (except one) were right-handed and gave their

² If attention is deployed according to the shape of the stimulus, attention would be allocated in a rectangular band instead of an annulus. However, ‘annulus’ is used to facilitate comparisons with results of other studies.

informed consent to participate in the study, which was approved by the local Ethical Committee. One participant was excluded from subsequent analyses, since his mean reaction time (RT) was more than two standard deviations above the mean of the group (in combination with the highest error percentage of the group).

Design

Each trial began with the presentation (1000 ms) of an outlined rectangle. The size of this rectangle cued the subject to attend the letter with an identical size in the three-level compound stimulus, which was subsequently presented for 450 ms on the position indicated by the cue. The compound stimulus consisted of three superimposed grey letters in three different sizes. Whereas the large (L) and the small (S) letter had fixed sizes ($7.0^\circ \times 3.0^\circ$ and $1.75^\circ \times 0.75^\circ$, respectively), the size of the middle letter could be large ($5.25^\circ \times 2.25^\circ$) or small ($3.5^\circ \times 1.5^\circ$), which defined the stimulus as ‘laM stimulus’ or ‘smM stimulus’, respectively. Note that the sizes of the levels of the smM stimulus are identical to the sizes used in experiment 2 (small stimulus) and 3. An example of this stimulus is shown in Figure 1. During the succeeding intertrial interval (2400 ms), in which a fixation cross was presented; participants had 2000 ms to report the identity of the letter (which could be a B, N, H, or S) at the cued level of the compound stimulus. Each of the three levels was cued equally often. Half of the participants responded to the letter B and N with the right index finger (key ‘5’ on the keypad of a keyboard) and to the letters H and S with the right middle finger (key ‘6’), whereas the stimulus-response coupling was reversed for the other half of the participants. After this 2000 ms period, feedback was given. A high tone (1000 Hz) indicated a correct response; a low tone (500 Hz) an erroneous response and a “buzz” tone was presented when no response was recorded.

Response congruency between the three levels was varied over trials. The letters on the three levels could all be coupled to the same response (congruent condition), or one letter corresponded to another response than the other two letters. The latter case constituted three types of incongruent conditions, according to the level at which the incongruent letter was presented (i.e., incongruent small level, incongruent middle level, or incongruent large level). All stimuli were presented in the middle of the screen on a grey background. Conditions were randomized with the constraint that each condition occurred equally often. Index and middle finger responses were randomized but occurred equally often, whereas presented letters were randomly chosen from the set that was associated with the selected responses.

The behavioral experiment consisted of three phases: First, stimulus-response coupling was trained. Then, participants were familiarized with the task (24 trials). Finally, each participant performed four runs of the main task, resulting in a total of 384 trials per subject. Participants could shortly rest between runs, leading to a total duration of about 40 minutes for each session. Throughout the experiment, participants were reminded to fixate the fixation cross, especially

during stimulus presentation, and to respond as fast and accurately as possible. Each participant was tested in a dimly-lit room while seated in front of a 19" LCD monitor with his or her head resting on a chin rest. The centre of the screen was at eye level and distance between the eyes and the monitor was 57 cm. Stimuli were presented and responses were recorded using the Presentation software package (Neurobehavioral Systems, San Francisco, USA).

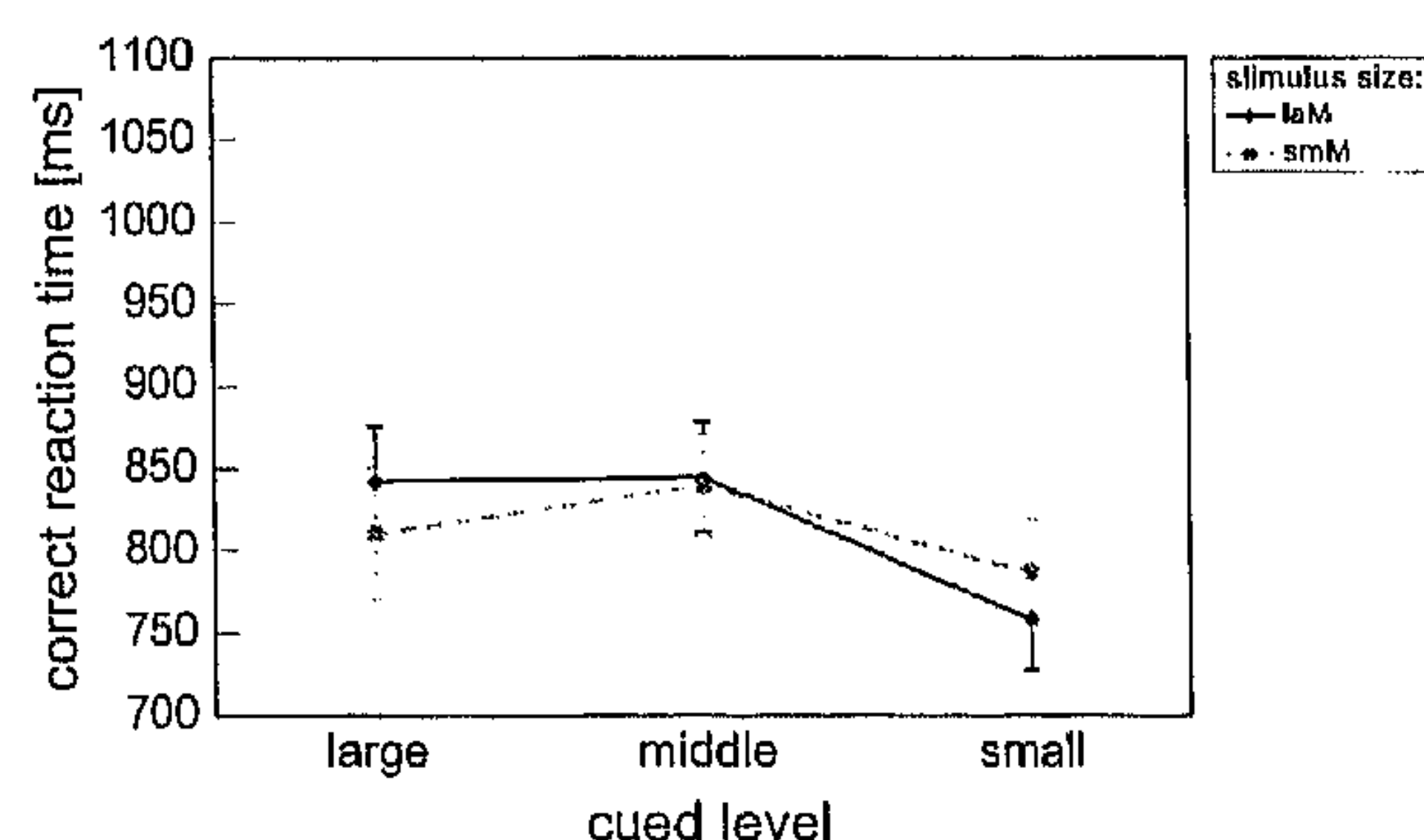
Data analysis

Only correct responses that occurred within a reaction time (RT) window defined as the interval from three standard deviations below to three standard deviations above the subject-specific mean reaction time per condition were included in the analyses of the present and all subsequent experiments. In none of the experiments, this led to a data loss of more than 1%. Mean RT of correct responses and mean error percentages of each participant were submitted to separate three-way repeated measures analyses of variance (ANOVAs) with stimulus size (laM stimulus, smM stimulus), cued level (attending the large, middle, or small level), and response congruency (all levels congruent, incongruent small level, incongruent middle level, incongruent large level) as factors. Greenhouse-Geisser correction was applied for p-values of these and all further results. The alpha-level was Bonferroni corrected for multiple post-hoc paired t-tests. In addition, all t-tests contrasting incongruent and congruent stimulus levels were one-sided, since only response congruency costs and not benefits were expected. For clarity, such t-tests are labeled 'one-sided' in the result section. Note that data of subsequent experiments are analyzed in a similar manner, unless stated otherwise.

Results

On average, participants responded 814(standard error: ± 32) ms after stimulus onset, with a mean accuracy of 93.8(± 1.4)%. As shown in Figure 2 and Table 1, a significant interaction between stimulus size and cued level was present in the RTs.

Figure 2. Mean reaction times of correct responses as a function of cued level and stimulus size. Error bars indicate standard errors of the mean (SEM).



Post-hoc paired t-tests suggested that in the laM stimulus, reaction times were slower for attending the large ($t(10) = 4.0$; $p = .002$) or the middle level ($t(10) = 4.8$; $p = .001$) compared to attending the small level. In the smM stimulus, attending the middle (but not the large) level was slower compared to attending the small level ($t(10) = 5.2$; $p = .0004$). The differences between the smM and laM stimulus, when the large or small level was cued, were not significant (large level: $p = .074$; small level: $p = .043$, which did not exceed the Bonferroni corrected alpha-level of 0.017).

Factors	RT		Error	
	F (df)	p	F (df)	p
Size	0.2 (1,10)	-	9.2 (1,10)	0.013*
Cued Level	11.9 (2,20)	0.003***	6.0 (2,20)	0.025*
RC	0.6 (3,30)	-	4.8 (3,30)	0.033*
Size • Cued Level	5.8 (2,20)	0.018*	5.5 (2,20)	0.014*
Size • RC	0.6 (3,30)	-	3.2 (3,30)	0.043*
Cued Level • RC	1.2 (6,60)	-	6.2 (6,60)	0.003***
Size • Cued Level • RC	2.6 (6,60)	-	4.9 (6,60)	0.004***

Table 1. F-values, degrees of freedom (df), and Greenhouse-Geisser corrected p-values for the main and interaction effects of the mean correct reaction time (RT; left columns) and error percentage (Error; right columns) analyses. Abbreviations: Size: stimulus size; RC: response congruency. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.005$.

All main and interaction effects in the error percentage analysis turned out to be significant (Table 1 & Figure 3). The presence of a significant three-way interaction motivated independent testing of each factor's level at each of the levels of the other two factors.

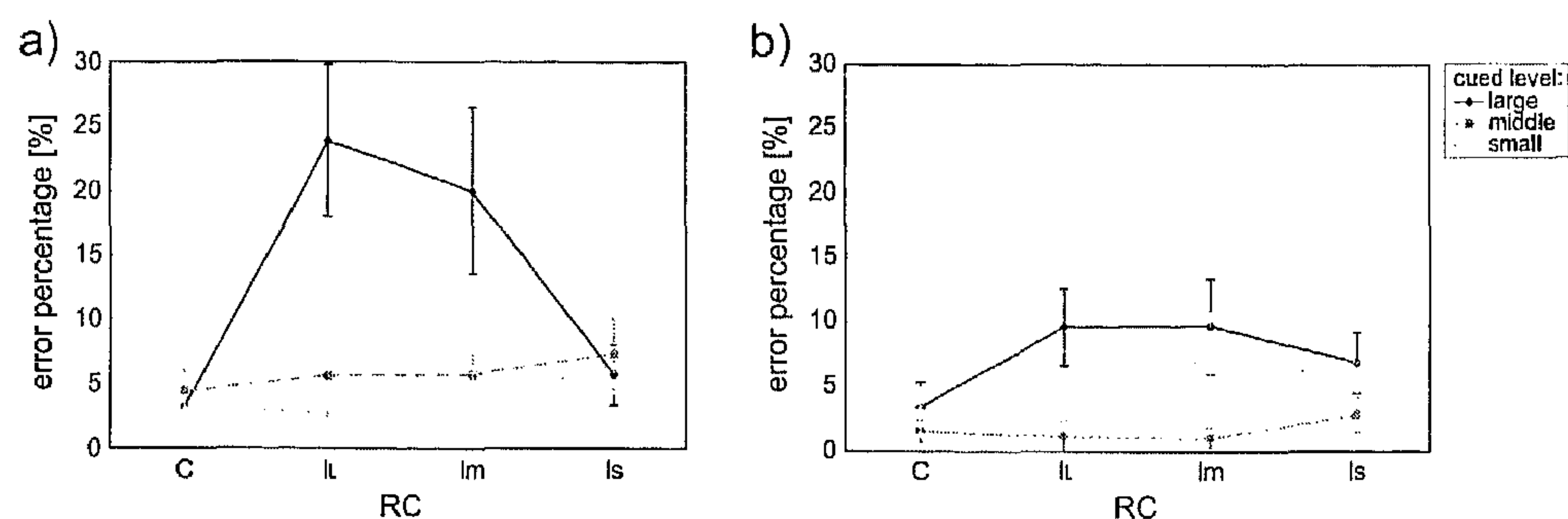


Figure 3. Mean error percentages as a function of response congruency type (C: all levels congruent, Il: incongruent large level, Im: incongruent middle level, Is: incongruent small level) and cued level separately shown for the a) laM stimulus and b) smM stimulus. Error bars indicate SEM.

To minimize the number of multiple comparisons, post-hoc analyses were focused on the primary interest of the study, the influence of stimulus size and cued level on response congruency. Therefore, the response congruency cost (RCC) of each condition was computed by subtracting the mean error percentage in the 'all levels congruent' condition from the large, middle, and small level incongruent conditions for each level of stimulus size and of cued level. This resulted in $(2[\text{size}] \times 3[\text{level}] \times 3[\text{incongruent} - \text{congruent level}]) = 18$ RCCs. Additionally, the significance of each RCC was assessed by testing the difference between the specific incongruent and the 'all levels congruent' condition. When the laM stimulus was presented, rather high RCCs were observed when the large level was

cued and this level (mean RCC = 20.5%) or the middle level (mean RCC = 16.5%) was incongruent. However, due to Bonferroni correction the RCC only reached significance when the incongruent letter was presented at the large level ($t(10) = 3.8$; $p = .002$; one-sided), whereas an incongruent letter at the middle level caused a tendency towards a significant RCC ($t(10) = 2.8$; $p = .009$; one-sided). None of the other RCCs reached significance.

Discussion

Response congruency did not affect mean RT in the laM nor in the smM stimulus. However, the size of the middle letter did influence the overall RT of responses to the different cued levels. In the laM stimulus, letters at the small level were identified faster compared to both other levels, whereas in the smM stimulus the mean RT for the small level was only faster than the middle but not the large level. The faster processing of the small compared to the other levels could have been caused by (a combination of) different factors. First, the result is in line with the predicted trade-off between the size of the zoom-lens and processing efficiency observed in other studies (e.g., Eriksen & St. James, 1986). Alternatively, faster processing of the small level could have been induced by a similar effect that causes faster processing of the local compared to global level in global-local attention studies, in which large compound stimuli (more than 8° ; Kinchla & Wolfe, 1979) or stimuli with relatively few local elements (Martin, 1979) were centrally presented with spatial certainty³. Furthermore, allocating attention in an annular (to solely attend the large or middle level) compared to circular (to attend the small level) fashion could have taken more time. Finally, the letter at the small level was the only letter not occluded by other letters, which could have contributed to its faster processing. To rule out that the faster processing of the small letter was not due to occlusion confounds, a second experiment was performed in which none of the letters was occluded (see experiment 4).

Error percentages were influenced by all factors (and possible interactions between the factors) in the design, complicating interpretations of the results. Since RCCs were of primary interest in the present experiment, the focus was on the analysis of these costs. Responses to the small letter were not influenced by the middle level, irrespective of the size of the middle level, indicating that the zoom-lens could be accurately scaled to encompass exclusively the small level. Likewise, no interference of the small or large level was found when the middle level was attended, suggesting that the zoom-lens could take on an annular shape that solely covered the middle level. In contrast, RCCs were obtained when the large level was cued, but only when the laM stimulus was presented. More

³ Note that several studies suggested that the global advantage effect (i.e., faster responses to the global than to the local level) observed by Navon (1977) depended on factors like the spatial uncertainty and the retinal position of the presented stimulus (Grice, Canham, & Boroughs, 1983; Lamb & Robertson, 1988).

precisely, significant RCC were present when the large level of the laM stimulus was incongruent, whereas an incongruent middle level tended to cause a RCC. No RCCs were obtained when the smM stimulus was presented, suggesting that attention could be accurately scaled to the large level if the distance between the large and middle level was large enough.

In sum, results suggested that the size and shape of the zoom-lens could fairly well be voluntarily controlled in this paradigm. The precise scaling to the small level, without interference of the surrounding, agrees with previous findings that the size of the zoom-lens can be contracted to cover a region with a diameter as small as 1° (Eriksen & Eriksen, 1974), in order to avoid interference of adjacent stimuli. Moreover, attention could be distributed in an annular shape, as suggested by the lack of interference of the other levels, when the middle level was attended. Finally, attention could also be precisely scaled (in an annular configuration) to the large level, but only when the distance between the large and middle level was large enough. Probably, the decrease in attentional resolution caused by an increase in eccentricity (e.g., Intriligator & Cavanagh, 2001) and the larger area that was covered by the zoom-lens, played a role in the less accurate scaling of attention to the large compared to the middle level. In the subsequent experiments, only the stimulus with the small (smM) and not the large (laM) size of the middle level was used. Therefore, based on the present results, accurate scaling to each of the levels of this stimulus was expected in the next experiments.

Experiment 2: continuous or discrete scaling of attention?

The previous experiment suggested that the zoom-lens could be precisely scaled to each level of the three-level ‘small middle level’ compound stimulus without experiencing interference of the other levels. In experiment 2, this stimulus was used to gain more insights in the mechanisms involved in scaling the zoom-lens. Previous studies investigating the scaling of the zoom-lens used a visual search paradigm in which the cue indicating the area in which the target could appear, varied in size (Greenwood & Parasuraman, 1999; Müller et al., 2003). However, before attention could be scaled to encompass the cued area, attention first needed to be shifted to this area, which led to results that were influenced by shifting as well as scaling processes. In contrast, the present experiment used a paradigm that allowed investigation of attentional scaling in relative isolation of other attentional processes. Attention was cued to the large or small level of the three-level compound stimulus. One-third of the trials were “zoom” trials, in which a letter-like symbol (the middle left figure in Figure 5) at the cued level indicated that attention had to be redirected to the opposite level (i.e., from large to small or v.v.) within the same compound stimulus. In this way, the rescaling of the zoom-lens could be studied separately from other processes like shifting attention (since the

cued regions were concentric, the attentional focus did not have to be moved during rescaling) and maintaining the established zoom-lens setting (immediately after rescaling, the target letter could be processed). Moreover, the use of a stimulus containing three levels allowed exploring whether the intermediate level can be “skipped” when attention is rescaled from the large to the small level or vice versa. The zoom-lens would temporarily encompass the intermediate letter, if the zoom-lens were rescaled continuously between the large and small level, but not if the zoom-lens could “jump” between the large and small level. The allocation of attention to the middle level was investigated by varying the response congruency between the letter of the middle and target level. RCCs due to incongruency between the middle and target level were expected if the zoom-lens were continuously rescaled, but not if the zoom-lens were discretely rescaled. Finally, the present experiment studied whether the degree of change in the zoom-lens’ size influenced rescaling processes, by including, next to the smM stimulus of experiment 1, a stimulus that was twice as large. If attention would be continuously rescaled with a constant velocity, more time would be required to accomplish rescaling between the large and small level in the large compared to small stimulus, since the change in the size of the zoom-lens would be larger in the large stimulus.

Methods

Participants

Fourteen healthy volunteers (4 males; mean age: 23.9 years, standard error: 1.7 years) with normal or corrected-to-normal visual acuity were paid for participation in this study. All participants were right-handed and none of them took part in the previous experiment. They gave their informed consent to participate in the study, which was approved by the local Ethical Committee. One participant was excluded from subsequent analyses, since his mean error percentage was more than two standard deviations above the group mean. In addition, eye movement recordings of 4 participants were lost due to file corruption.

Design

The design was similar to that of experiment 1, unless stated otherwise. Each trial began with the presentation (500 ms) of an outlined rectangle of which the size (and the position) matched the size of either the large or the small character of the three-level compound stimulus that was subsequently presented (300 ms) after a variable cue delay (500 ± 200 ms). There were two sizes of the compound stimulus that consisted of a superimposed large, middle, and small grey character. The sizes of the levels in the small compound stimulus (Figure 1) were $7.0^\circ \times 3.0^\circ$ (large), $3.5^\circ \times 1.5^\circ$ (middle), and $1.75^\circ \times 0.75^\circ$ (small), whereas the large compound stimulus sizes were $14.0^\circ \times 6.0^\circ$ (large), $7.0^\circ \times 3.0^\circ$ (middle), and $3.5^\circ \times 1.5^\circ$

(small), respectively. Note that each level was half the size of the larger level and that each level of the large stimulus was twice as large as the corresponding level of the small stimulus. Similar to experiment 1, if a letter (which could be a B, N, H, or S) was presented at the cued level of the compound stimulus, the identity of this letter should be reported during the succeeding intertrial interval (2200 ± 200 ms), after which auditory feedback was given. In contrast to the trials on which participants could *stay* attending the cued level, participants were required to rescale their attention to the opposite level when the zoom symbol was presented at the cued level (33% of the trials). That is, if the large level was cued and a zoom symbol was presented at this level, participants were required to ‘*zoom-in*’ their spatial attention window and report the identity of the letter presented at the small level. Likewise, if a zoom symbol was presented on the cued small level, participants should ‘*zoom-out*’ their spatial attention window and report the identity of the large letter. These four types of trials resulted in four conditions, labeled *stay-large* and *stay-small* when attention did not have to be rescaled to another level (*stay* trials) and *zoom-in* and *zoom-out* if attention had to be rescaled to the opposite level (*zoom* trials). Similar to the first experiment, the letters on the three levels in *stay* trials were all coupled to the same response, or one letter corresponded to another response than the letters at the other two levels. In the *zoom* conditions on the other hand, the compound stimulus contained two letters (at the middle and target level) and a neutral (i.e., not coupled to a response) zoom symbol (at the cued level). These two letters could be response congruent or incongruent, leading to two response congruency variants for the compound stimulus in *zoom* trials.

The zoom symbol was never presented at an uncued level and conditions were counterbalanced according to a Latin square design. The experiment consisted of three phases: First, stimulus-response coupling was trained. Then, participants were familiarized with the task, first by using a version of the task in which every phase of the trial lasted twice as long compared to the main task, followed by training the task with a trial timing identical to the main task. Finally, each participant performed four runs of the main task (576 trials) in which the movement of the left eye was monitored using an Eyelink tracker with a sampling rate of 250 Hz (SensoMotoric Instruments GmbH, Teltow, Germany). Eye movements were categorized as saccades when the movement velocity exceeded $35^\circ/\text{s}$ or when the movement acceleration was higher than $9500^\circ/\text{s}^2$. Participants were instructed to fixate the fixation cross (especially during stimulus presentation) and recorded eye movements were monitored online by the experimenter. Between runs participants were reminded that maintaining fixation was important and participants were notified if they did not fixate properly.

Data analysis

Data analysis was similar to that of the previous experiment, except for the following. Mean RTs and mean error percentages of stay trials were submitted to separate three-way ANOVAs with stimulus size (large, small), cued level (attending the large or small level), and response congruency (all levels congruent, incongruent small level, incongruent middle level, incongruent large level) as factors. In addition, an ANOVA including stimulus size (large, small), rescaling (from large to small, from small to large), and response congruency (congruent, incongruent) as factors, was performed on RTs and error percentages obtained in zoom trials.

Results

Although responses in zoom (mean (\pm standard error): 1162 (\pm 43) ms) trials were much slower than in stay (900 (\pm 42) ms) trials ($t(12) = 16.2$; $p = .000000002$), accuracy did not differ between zoom (90.6(\pm 1.1)%) and stay (92.1(\pm 1.0)%) trials ($t(12) = 1.5$; $p = .168$). In addition, eye movement analyses revealed that the average number of saccades ($t(8) = 0.8$; $p = .470$) and the average deviation of the eyes from the fixation cross ($t(8) = 0.9$; $p = .410$) during stimulus presentation, did not differ between zoom and stay trials.

None of the (interactions between) factors influenced mean RT in stay or in zoom trials (Table 2). In contrast, analysis of the error percentages revealed an effect of response congruency on stay trials (Table 2a). Post-hoc comparisons showed that mean error percentage was higher when the large ($t(12) = 5.7$; $p = .00005$; one-sided) or small ($t(12) = 4.2$; $p = .0006$; one-sided) level was incongruent compared to congruent. All interactions of the response congruency factor with the other factors showed a trend towards significance ($p < .09$) in stay trials. Since this suggested that the other factors might have influenced congruency, the RCC of each condition was computed and tested (see experiment 1). When the large level was attended in the small stimulus, significant RCCs were present when the large (mean RCC = 15.1%; $t(12) = 3.9$; $p = .001$; one-sided) or small (mean RCC = 11.9%; $t(12) = 4.3$; $p = .0005$; one-sided) level was incongruent. Likewise, an incongruent large level could elicit an RCC if this level was attended in the large stimulus (mean RCC = 7.7%; $t(12) = 3.8$; $p = .0015$; one-sided). However, in contrast to the small stimulus, the RCC of the small incongruent level when the large level was attended did not survive Bonferroni correction in the large stimulus (mean RCC = 4.5%; $t(12) = 2.5$; $p = .014$; one-sided).

a) stay-trials

Factors	RT		Error	
	F (df)	p	F (df)	p
Size	2.9 (1,12)	-	1.8 (1,12)	-
Cued Level	1.8 (1,12)	-	1.4 (1,12)	-
RC	0.7 (3,36)	-	17.2 (3,36)	0.000***
Size • Cued Level	0.1 (1,12)	-	2.0 (1,12)	-
Size • RC	1.0 (3,36)	-	2.7 (3,36)	-
Cued Level • RC	1.1 (3,36)	-	2.6 (3,36)	-
Size • Cued Level • RC	2.2 (3,36)	-	3.1 (3,36)	-

b) zoom-trials

Factors	RT		Error	
	F (df)	p	F (df)	p
Size	0.2 (1,12)	-	1.0 (1,12)	-
Rescaling	0.0 (1,12)	-	7.0 (1,12)	0.021*
RC	4.3 (1,12)	-	18.2 (1,12)	0.001***
Size • Rescaling	0.3 (1,12)	-	8.7 (1,12)	0.012*
Size • RC	3.0 (1,12)	-	1.4 (1,12)	-
Rescaling • RC	0.0 (1,12)	-	6.8 (1,12)	0.026*
Size • Rescaling • RC	0.1 (1,12)	-	1.6 (1,12)	-

Table 2. F-values, degrees of freedom (df), and Greenhouse-Geisser corrected p-values for the main and interaction effects of the mean correct reaction time (RT; left columns) and error percentage (Error; right columns) analyses of a) stay and b) zoom trials. Abbreviations: Size: stimulus size; RC: response congruency. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.005$.

In zoom trials, the stimulus size and rescaling factors interacted in the error analysis (Table 2b and Figure 4a). Post-hoc comparisons showed that more errors were made when attention was zoomed-in compared to zoomed-out, but only in the large stimulus ($t(12) = 3.9$; $p = .002$). A tendency towards more errors in the small compared to the large stimulus for zooming-out ($t(12) = 2.4$; $p = .034$) appeared to contribute to this influence of stimulus size on zooming-in versus zooming-out. In addition, the rescaling and response congruency factors interacted (Table 2b and Figure 4b). The error RCC for an incongruent middle and target level was computed for zoom-in and zoom-out trials and the significance of the RCC was assessed by testing the difference between the incongruent and congruent stimulus in each condition. This resulted in significant RCCs for both zoom-in (RCC: 9.6%; $t(12) = 4.1$; $p = .0005$; one-sided) as well as zoom-out (RCC: 3.0%; $t(12) = 2.1$; $p = .030$; one-sided) trials⁴. However, the RCC for zooming-in was higher than for zooming-out (mean difference = 6.6%; $t(12) = 2.6$; $p = .023$).

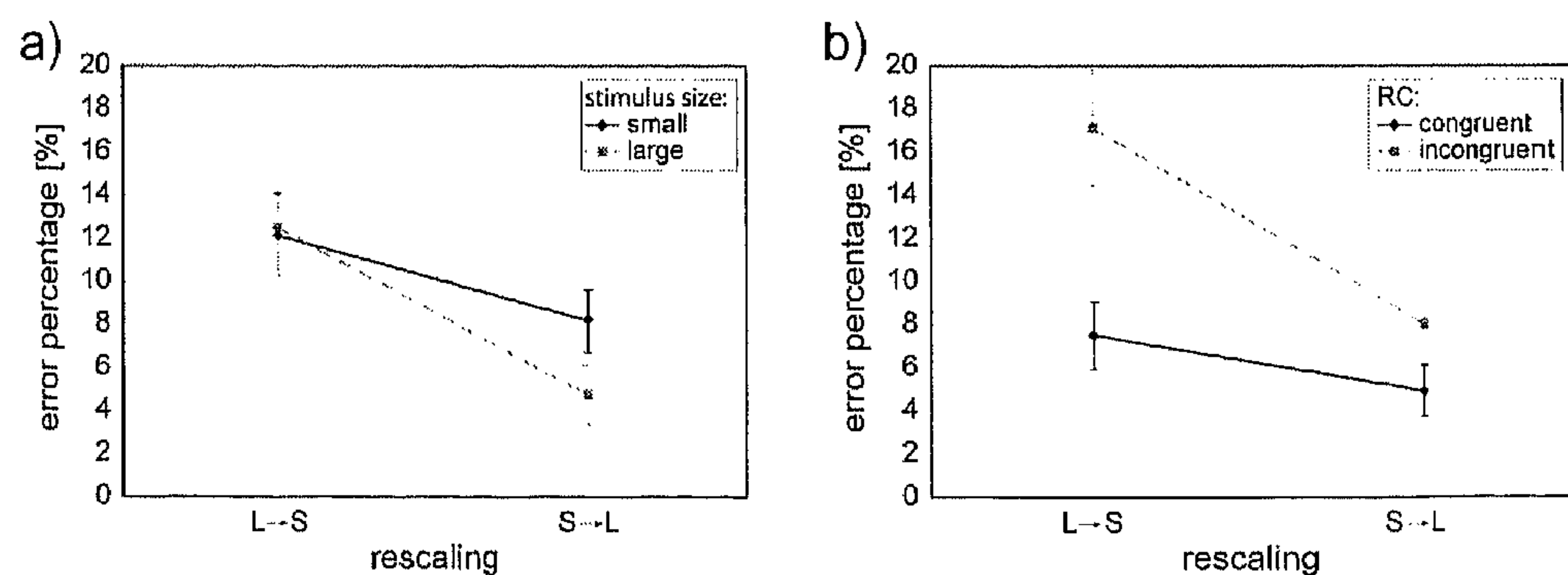


Figure 4. Mean error percentages as a function of rescaling (L→S: zoom-in from large to small level; S→L: zoom-out from small to large level) and a) stimulus size or b) response congruency type. Error bars indicate SEM.

⁴ To compare rescaling from the small to the large level in this and the next experiment, the RCC of zoom-out was also separately tested for the small stimulus (RCC: 3.5%; $t(12) = 2.2$; $p = .025$; one-sided).

Discussion

In line with results of the small stimulus presented in experiment 1, responses to the large and small level were equally fast in stay trials. However, in contrast to the previous experiment, more errors were made in stay trials when the small or large level was incongruent with the other levels. This was surprising since this effect (or a tendency towards this effect) was not obtained in experiment 1 and, moreover, was also not observed in the experiment described in Chapter 5 that used an identical task (except for trial timing and the use of only the small stimulus size). To explore this unexpected finding in more detail, error RCCs of each condition were analyzed. An incongruent compared to congruent large level caused more errors when this level was cued in both the small and large stimulus. In addition, a significant error RCC was elicited when the large level was attended in the small stimulus in which the small level was incongruent. These results suggested that when the large level was attended, some attention was also deployed to the small level. Possibly, the strong emphasize on the importance of maintaining fixation (in combination with the continuous eye-movements monitoring) could have motivated participants to develop a strategy in which the fixation cross was attended even when the large level was cued. As suggested by the enhanced activations in retinotopic areas representing the region in the visual field surrounding the fixation cross when this is attended (Chapter 7), attending the fixation cross could have elicited an attentional distribution that encompassed not only the fixation cross location but also its surroundings, thereby covering parts of the small letter. Moreover, accumulating evidence (e.g., McMains & Somers, 2004; Chapter 7; for review, see Chapter 6) suggests that attention can be simultaneously allocated to two non-contiguous regions. However, since dividing attention appears to entail an increase in task load (Nebel et al., 2005), attention might only be divided if prompted by task demands, such as in the present experiment.

The primary research question of this experiment was related to the rescaling of the zoom-lens in the so-called zoom-trials. To explore whether attention was rescaled in a continuous or discrete fashion, responses to stimuli representing a response incongruent versus congruent middle and target level were contrasted. A significant RCC was observed both when attention was zoomed-in from the large to the small level, as well as when attention was zoomed-out from the small to the large level. These results indicated that the zoom-lens temporarily encompassed the intermediate letter during rescaling. Since RCCs caused by interference of the intermediate level could not be avoided, the zoom-lens apparently can not be discretely rescaled between the small and large level. Interestingly, higher error RCCs were obtained for zooming-in compared to zooming-out. This indicated that, although the middle level interfered irrespective of whether the zoom-lens is contracted or expanded, the degree of interference depended on the “direction” in which the zoom-lens was rescaled. Given that

attention seems to be able to influence scene segmentation processes (Driver et al., 2001; Vecera, Fievaris, & Filapek, 2004; Zenger, Braun, & Koch, 2000), zooming-in and zooming-out might have had different influences on the figure-ground segmentations of the three levels. For example, due to an increasing number of overlapping objects towards the middle of the stimulus (see Figure 1), it is more difficult to segment the small letter from its background when zooming-in than segmenting the large letter from its background when zooming-out, which increases the possibility of interference from the middle level for zooming-in compared to zooming-out. Therefore, the use of this specific stimulus might have influenced the divergent findings for contracting or expanding the zoom-lens. In addition to exploring the effect of the direction of rescaling, the influence of the extent of rescaling was studied by including two stimulus sizes in the present experiment. However, no indication was found for slower responses when the zoom-lens had to be rescaled to a greater extent. In fact, the RT analyses revealed no effect at all. The only obtained effect of size was the unexpected tendency towards fewer errors for zooming-out in the large compared to small stimulus, which might have been caused by differences in segmentation difficulty between the two stimulus sizes (i.e., the large letter is easier to segment from its background in the large compared to small stimulus). Therefore, the influence of the extent of rescaling was investigated in more detail in experiment 3, in which only one stimulus size was used. Finally, eye-movements were recorded throughout the experiment, to control that rescaling of the zoom-lens was not accompanied by overt shifts of attention. Since the letters of the stimulus were concentric, the center of the attentional focus did not have to be shifted during rescaling. Therefore, we did not expect a higher number of saccades or stronger deviation from the fixation cross for zoom compared to stay trials, which was confirmed by the performed eye-movements analysis.

Overall, this experiment showed some intriguing results. On stay trials, the small level interfered when attention was cued to the large level, which was probably due to attention to the fixation cross. However, when attention had to be rescaled from the large to the small level, the intermediate level strongly interfered. This suggested that attention cannot be discretely rescaled between the large and middle level, even though findings in stay trials suggest that when the large level is attended, some attention is also deployed to the small level. Possibly, more attention is needed to identify the small letter for an accurate response than the amount of attention that was allocated to the fixation cross and its surroundings in the stay trials. However, the most important finding was that performance was deteriorated when attention was zoomed-in or zoomed-out in an incongruent compared to a congruent stimulus. This RCC indicated that the zoom-lens temporarily encompassed the intermediate letter during rescaling, thereby suggesting that the attentional zoom-lens can not be rescaled in a discrete fashion.

Experiment 3: are attentional scaling mechanisms influenced by the direction and the extent of scaling?

Results of experiment 2 suggested that the zoom-lens could not be discretely contracted or expanded. However, the extent of rescaling did not appear to influence RT, which would be expected if attention would be continuously rescaled with a constant velocity. Experiment 3 investigated the influence of the direction and extent of rescaling on rescaling mechanisms in more detail, by extending the rescaling variants of experiment 2 with conditions in which attention had to be rescaled from and to the middle level.

Methods

Participants

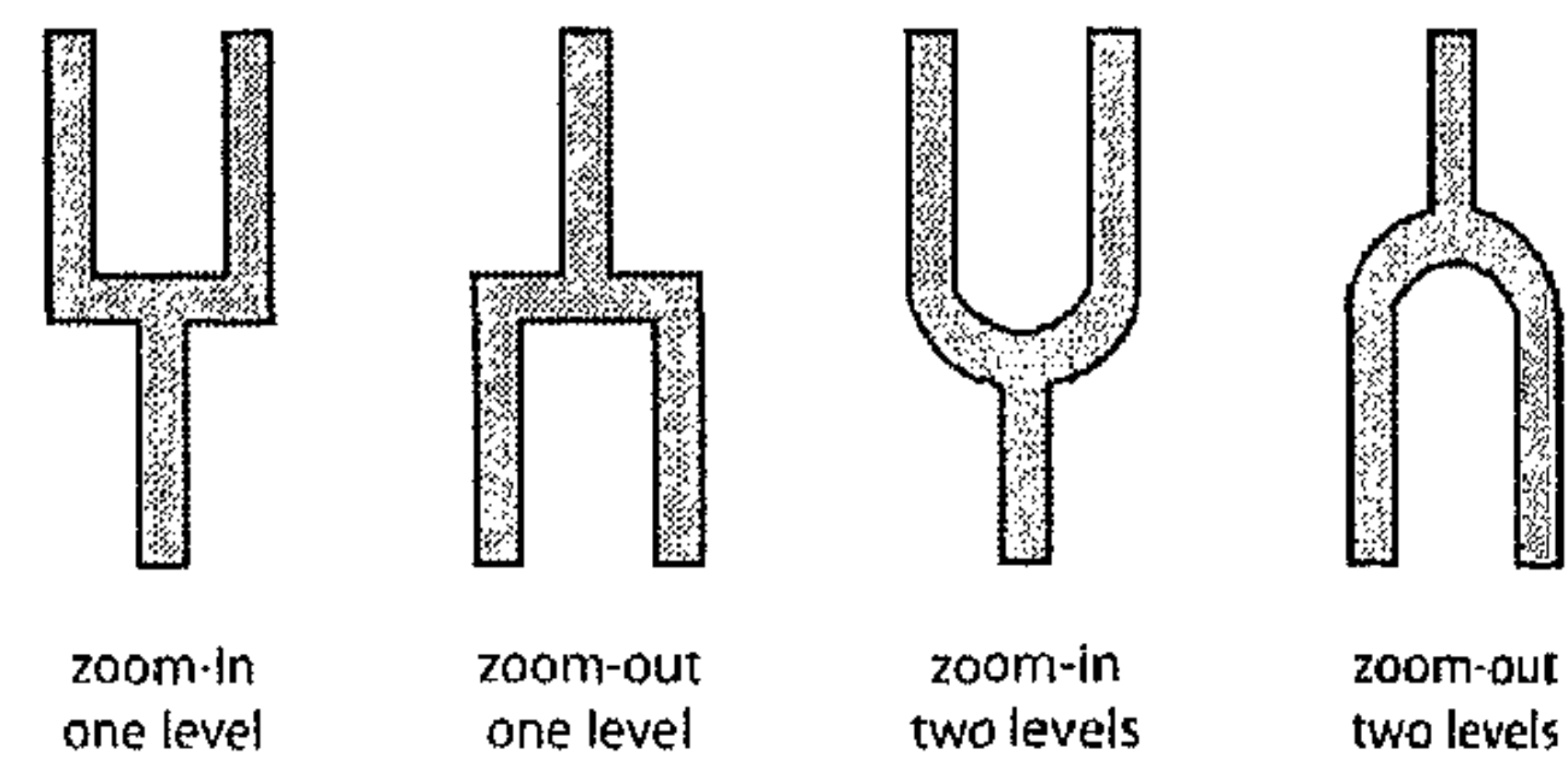
Twelve healthy volunteers (3 males; mean age: 23.2 years; standard error: 2.5 years) with normal or corrected-to-normal visual acuity were paid for participation in this study. All participants were right-handed and none of them took part in the previous experiments. They gave their informed consent to participate in the study, which was approved by the local Ethical Committee. One participant, having a mean error percentage of two standard deviations above the group mean, was excluded from subsequent analyses.

Design

The design of the experiment was similar to the previous experiment, except for the following modifications. Only the small stimulus of the previous experiment was used. The durations of the cue-delay (600 ± 200 ms), compound stimulus presentation (450 ms), and the intertrial interval (2400 ± 200 ms) slightly deviated from those in experiment 2. Similar to the previous experiment, the current design contained both *stay* (50%) and *zoom* (50%) trials. In contrast to the previous experiment, the middle level could be cued (in both types of trials) and, moreover, there were four variants of the zoom symbol (Figure 5) that were presented at the cued level in zoom trials. This zoom symbol instructed participants in which direction (i.e., zoom-in or zoom-out) and to which level (i.e., rescale one or two levels) attention had to be rescaled. To illustrate, if the outer left figure was presented on the middle level, attention had to be “zoomed-in” to the small level. In contrast, if the outer right figure was presented on the small level, attention had to be “zoomed-out” to the large level. Note that the zoom symbol was never presented at an uncued level. Conditions were randomized, but each condition occurred equally often. The behavioral experiment consisted of several phases: First, stimulus-response coupling was trained. Then, participants were familiarized with the stay trials and subsequently with the zoom trials, in a version of the task in which the stimulus was presented for 900 ms. Subsequently, stay and zoom

trials were trained intermixed, first with a stimulus presentation of 900 ms and subsequently with 450 ms presentation. Finally, each participant performed seven runs of the main task, resulting in a total of 672 trials per subject. Participants could rest between runs, leading to a total duration of about 90 minutes for each session. A 17" CRT monitor was used to present the stimuli, whereas all other settings were similar to the previous experiments.

Figure 5. The zoom symbols presented on the cued level of the compound stimulus in “zoom” trials. The comment below each symbol indicates to which level (rescale to adjacent level [left figures] or skip adjacent level and rescale to the subsequent level [right figures]) and in which direction (zoom-in to a smaller level or zoom-out to a larger level) attention had to be rescaled. Note that the middle right figure could therefore only be presented on the large level (rescale to small level) and the outer right figure only at the small level (rescale to large level). The second symbol from left was also used in zoom trials of experiment 2.



Data analysis

Data analysis was similar to experiment 2, except that stimulus size was not included as a factor in the two-way ANOVAs. In addition, the cued level factor contained three (small, middle, and large level) and the rescaling factor contained six (from large to middle, from middle to small, from large to small, from small to middle, from middle to large, and from small to large) instead of two levels. To directly test the effects of extent and direction of rescaling, and their possible interaction, zoom-trial data were submitted to a second ANOVA with direction and extent as factors. The direction factor contained a zoom-in (from large to middle, from middle to small, from large to small) and zoom-out (from small to middle, from middle to large, from small to large) level, whereas the two levels of the extent factor were an one-level (i.e., trials in which attention was rescaled to an adjacent level: from large to middle, from middle to small, from small to middle, from middle to large) and a two-level (i.e., rescaling of attention to the opposite level: from large to small, from small to large) cluster. The RCCs obtained in zoom-trials were also submitted to this latter ANOVA to study the influence of the extent and direction of rescaling on response congruency effects.

Results

On average, participants were slower ($t(10) = 13.0$; $p = .0000001$) and less accurate ($t(10) = 5.2$; $p = .0004$) in zoom (RT: 1378 (± 53) ms; accuracy: 89.9(± 1.6)%) compared to stay (RT: 982 (± 40) ms; accuracy: 95.8(± 0.8)%) trials. As shown in Table 3a, a main effect of cued level and congruency was obtained in

the RT analyses of stay trials, whereas none of the effects became significant in the error analyses of the stay trials. Post-hoc comparisons revealed that responses to the middle ($t(10) = 4.4$; $p = .001$) or small ($t(10) = 2.92$; $p = .016$) level were slower than responses to the large level. Additional post-hoc comparisons indicated that the main effect of response congruency was caused by slower responses when the middle ($t(10) = 3.3$; $p = .004$; one-sided) or small ($t(10) = 3.8$; $p = .002$; one-sided) level was incongruent compared to congruent. The presence of a weak tendency towards interaction between the two factors ($F(6,60) = 2.4$; $p = .089$) in combination with significant RCCs for an incongruent cued middle level (mean RCC = 64.0 ms; $t(10) = 3.6$; $p = .003$; one-sided) or an incongruent cued small level (mean RCC = 44.4 ms; $t(10) = 4.3$; $p = .001$; one-sided) but no RCCs for the cued large level, suggested that response congruency effects could have contributed to the obtained slower responses to the middle and small level compared to the large level.

a) stay-trials

Factors	RT		Error	
	F (df)	p	F (df)	p
Cued Level	7.7 (2,20)	0.005**	0.0 (2,20)	-
RC	6.3 (3,30)	0.006**	2.6 (3,30)	-
Cued Level * RC	2.3 (6,60)	-	1.5 (6,60)	-

b) zoom-trials

Factors	RT		Error	
	F (df)	p	F (df)	p
Rescaling	2.3 (5,50)	-	5.4 (5,50)	0.001**
RC	10.8 (1,10)	0.008**	13.2 (1,10)	0.005**
Rescaling * RC	3.6 (5,50)	0.020*	1.2 (5,50)	-

Table 3. F-values, degrees of freedom (df), and Greenhouse-Geisser corrected p-values for the main and interaction effects of the mean correct reaction time (RT; left columns) and error percentage (Error; right columns) analyses of a) stay and b) zoom trials. RC: response congruency. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.005$.

In zoom trials (Table 3b), the rescaling and response congruency factors interacted in the RT analysis (Figure 6a), which motivated the computation and testing of each condition's RCC. Significant RCCs were present when rescaling the zoom-lens from the large to the middle level (mean RCC = 84.3 ms; $t(10) = 3.0$; $p = .006$; one-sided) or from the small to the middle level (mean RCC = 47.1 ms; $t(10) = 2.9$; $p = .008$; one-sided). The error analysis showed a main effect for both response congruency as well as rescaling, but the two factors did not interact (Table 3b & Figure 6b). As expected, the main effect of response congruency revealed a higher error percentage for incongruent compared to congruent trials ($t(10) = 3.6$; $p = .003$; one-sided). The main effect of rescaling was not explored by pair-wise testing all 15 rescaling combinations, which would inflate the probability of a Type I error, but by clustering rescaling levels according to their direction and extent of rescaling and perform an additional ANOVA on these groups as discussed below. Finally, although no interaction between rescaling and response congruency was present, error RCCs were tested as planned for rescaling attention from the large to the small level and v.v., to compare results of experiment 2 and 3. Unlike the previous experiment, interference of the middle level was only obtained for rescaling from the large to the small level (mean RCC = 10.1%; $t(10) = 1.9$; $p = .043$; one-sided), whereas no significant error RCCs

were obtained for rescaling from the small to the large level (mean RCC = 0.0009%; $t(10) = 0.0003$; $p = .500$; one-sided).

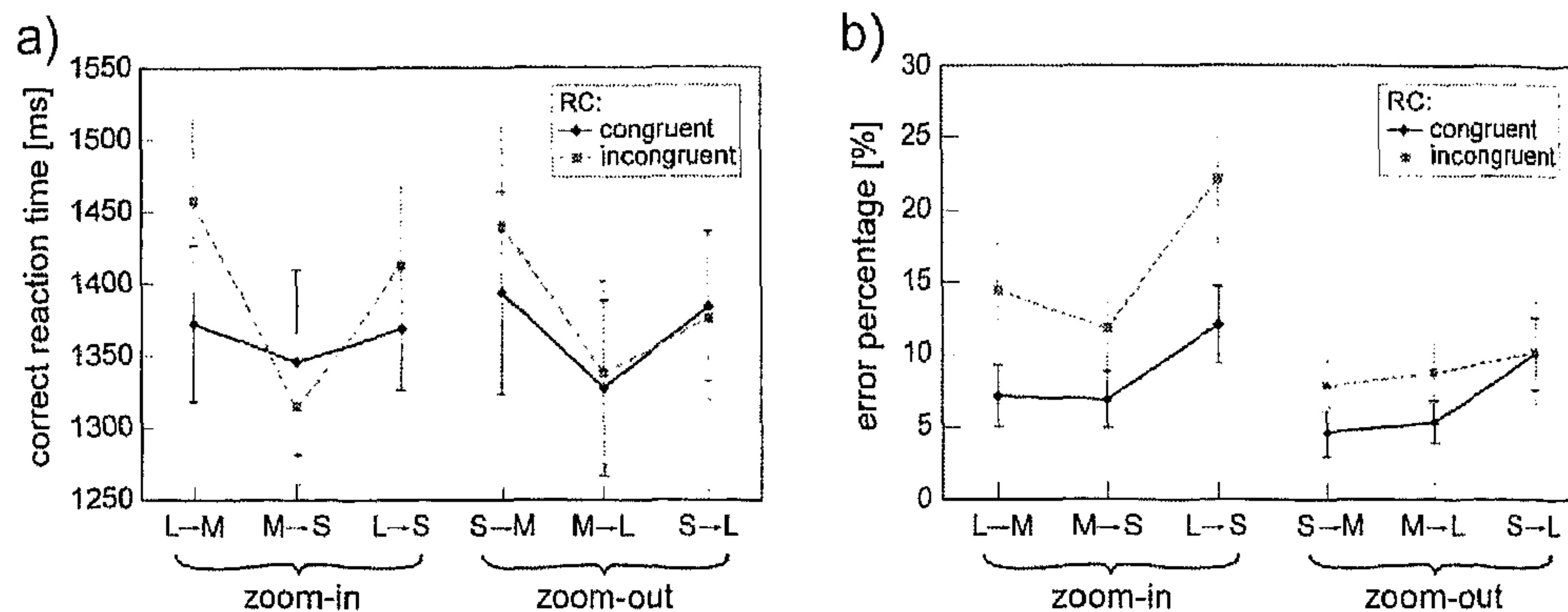


Figure 6. Mean a) correct reaction times and b) error percentages as a function of rescaling (L→M: zoom-in from large to middle level; M→S: zoom-in from middle to small level; L→S: zoom-in from large to small level; S→M: zoom-out from small to middle level; M→L: zoom-out from middle to large level; S→L: zoom-out from small to large level) and response congruency (RC) type. Error bars indicate SEM.

To directly test the influence of the direction and extent of rescaling, zoom-trial data were submitted to additional ANOVAs. No significant effects were obtained in the RT analysis ($p > .550$) or in the RT RCC analysis ($p > .150$). In contrast, main effects of direction ($F(1,10) = 13.6$; $p = .004$) and extent of rescaling ($F(1,10) = 9.4$; $p = .012$) but no interaction ($p = .294$) was obtained in the error analysis. As expected, a higher error percentage was revealed for zoom-in ($13.5(\pm 2.0)\%$) compared to zoom-out ($8.3(\pm 1.7)\%$) trials. Moreover, lower error percentages were obtained if attention could be rescaled to an adjacent ($8.3(\pm 1.6)\%$) instead of opposite ($13.6(\pm 2.3)\%$) level. In addition, analysis of the error RCCs revealed a main effect of direction ($F(1,10) = 5.3$; $p = .044$), indicating that higher errors RCCs were experienced when attention was zoomed-in (mean RCC = 8.0%) compared to zoomed-out (mean RCC = 1.7%). None of the other effects became significant.

Responses on zoom-trials were on average almost 400 ms slower than on stay-trials, which indicated that a considerable amount of time is needed to rescale attention. Since the 450 ms stimulus duration might have been too short to properly rescale attention we performed an additional experiment (with six new participants) using an identical design but with an increased stimulus duration (750 ms). In this experiment, which is subsequently referred to as experiment 3B (as opposed to experiment 3A, with 450 ms stimulus duration), none of the main effects or their interaction was significant in both the RT as well as error analyses of stay-trials. In contrast in the RT analysis of the zoom-trials, a main effect of rescaling ($F(5,25) = 4.3$; $p = .048$) and of response congruency ($F(1,5) = 9.2$; $p = .029$), as well as an interaction between these factors ($F(5,25) = 4.8$; $p = .029$), was

obtained. RCCs were computed for each condition to explore the interaction effect. However, in contrast to experiment 3A none of the RCCs became significant. Only a marginal trend was observed for zooming-in from the large to the small level (mean RCC = 47.1 ms; $t(5) = 2.9$; $p = .017$; one-sided). The error analysis of the zoom-trials revealed an interaction between rescaling and response congruency ($F(5,25) = 4.0$; $p = .032$), but the main effects did not become significant. Similar to experiment 3A, planned comparisons showed that interference of the middle level was in this experiment only obtained for rescaling from the large to the small level (mean RCC = 12.5%; $t(5) = 3.1$; $p = .003$; one-sided), but not for rescaling the other way around (mean RCC = -0.6%).

The influence of the direction and extent of rescaling on RT in zoom-trials was studied in an additional ANOVA that revealed slower responses for zooming-in compared to zooming-out ($F(1,5) = 11.2$; $p = .020$). Unexpectedly, rescaling to the opposite level (mean = 1362 ms) was not slower than to an adjacent level (mean = 1329 ms) ($F(1,5) = 2.9$; $p = .150$). When this difference was further tested conform to our hypothesis that a greater magnitude of change in size of the zoom-lens would take more time, a marginal trend was observed ($t(5) = 1.7$; $p = .075$; one-sided). The RT RCC analysis, on the other hand, showed a stronger effect of extent ($F(1,5) = 36.6$; $p = .002$), indicating that more interference was experienced when attention was rescaled to the opposite instead of adjacent level. The error analysis also revealed a main effect of the extent of rescaling ($F(1,5) = 7.1$; $p = .045$), resulting from more errors for rescaling to the opposite level compared to an adjacent level. Finally, analysis of the error RCCs revealed a main effect of direction ($F(1,5) = 10.4$; $p = .023$) as well as extent ($F(1,5) = 20.8$; $p = .006$), indicating that higher errors RCCs were associated with zooming-in compared to zooming-out and with rescaling to the opposite compared to an adjacent level, respectively. None of the other effects reached significance.

Discussion

Responses on trials in which attention was rescaled were on average almost 400 ms slower compared to trials in which attention did not need to be rescaled. This is in line with a study by Stoffer (1993), which suggested that voluntary scaling of attention is time-consuming, being about twice as slow as shifting attention. In the present and previous experiments, response congruency had a stronger effect on accuracy than on RT, which contrasts with findings in flanker paradigms (e.g., Eriksen & Eriksen, 1974). The considerable amount of time needed to rescale attention, in combination with the unexpected strong accuracy effects, could indicate that the stimulus duration was too short to properly rescale attention (in a subset of trials). Therefore, an additional experiment (experiment 3B) was performed with prolonged stimulus duration (750 ms), which however will only be discussed in the context of the main experiment (experiment 3A).

In stay trials of experiment 3A, incongruency of either the small or the middle level with the other two levels appeared to slow responses, especially when the small or middle level was attended. This suggested that attention was not solely scaled to the small or middle level, but that the zoom-lens encompassed both levels during stay trials. The slightly broader scaling of the (annulus' width of the) zoom-lens might have been a strategy to prepare for faster rescaling if a zoom-trial would appear. The disappearance of this effect when stimulus duration was prolonged, suggested that the broader scaling of the zoom-lens could indeed have been a strategy to rescale as fast as possible if needed, although this strategy did slow responses to the small and middle level when attention did not have to be rescaled. However, the primary focus was on the rescaling of the zoom-lens in zoom-trials. Unlike the previous experiment, response congruency effects were not only obtained in the error but also in the RT analyses of zoom-trials. Interestingly, an RT RCC was obtained when the zoom-lens had to be rescaled from the small to the middle level, indicating that incongruency between the large and middle level decreased performance even though the large level did not need to be attended. Moreover, a similar cost was obtained when attention was rescaled from the large to the middle level caused by incongruency between the small and middle level. These effects suggested that when attention was rescaled from the large or small level to the middle level, there was an "overshoot" in the extent of rescaling. That is, when attention was rescaled from the small to the middle level, the zoom-lens might have been rescaled a bit broader than the middle level, thereby encompassing the large level. Due to a similar overshoot, the zoom-lens could have encompassed the small letter when rescaled from the large to the middle level. In contrast, no RCCs were present when attention had to be rescaled to the middle level if stimulus duration was prolonged to 750 ms. Although the disappearance of the effect could have been due to the limited number of participants in experiment 3B, it could alternatively indicate that attention can be more precisely rescaled if more time is available. Furthermore, the error analyses of the zoom-trials in experiment 3A revealed that although the middle level interfered when attention was zoomed-in from the large to the small level, it did not interfere when attention was rescaled the other way around, in contrast to experiment 2. A similar result was obtained in experiment 3B. Experiment 2 and 3 (3A as well as 3B) contained about the same number of trials in which attention had to be rescaled from the small to the large level in the (small) stimulus (48 and 56 trials per participant, respectively), but the overall number of trials in which attention had to be zoomed-out (irrespective of the cued and target level and stimulus size) was 1.75 times larger in experiment 3A and 3B compared to experiment 2. The presence of only one instead of two stimulus sizes, in combination with the larger amount of zooming-out trials (leading to increased zooming-out experience), could have contributed to an optimization of zooming-out from the small to the large level in experiment 3A and 3B, thereby decreasing interference of the middle level. Instead of mutually testing the six rescaling levels

in zoom-trial analyses, the rescaling levels were clustered to explore the influence of the direction and extent of rescaling in more detail. Although neither the direction nor the extent of rescaling had influence on RT or RT RCCs, both factors had an effect on accuracy: higher error percentages were revealed for zooming-in than for zooming-out trials and for rescaling to the opposite (i.e., from large to small or v.v.) compared to an adjacent level. In addition, analysis of the error RCCs showed higher errors RCCs for zoom-in compared to zoom-out trials.

When stimulus duration was prolonged (experiment 3B), the effects on accuracy became less strong at the expense of an increase in RT effects. For example, zooming-in compared to zooming-out led to slower instead of less accurate performance. In addition, the difference in accuracy between rescaling to the opposite instead of an adjacent level became less strong (but was still significant), whereas this difference started to show a tendency towards significance in RT. However, this tendency was rather weak and therefore did not fully confirm the expectation that rescaling to the opposite instead of adjacent level would take more time. The interpretation of this result is not straightforward. On the one hand, there could be truly no influence of the extent of rescaling on the time needed for rescaling, as suggested by the results of this and the previous experiment. This would imply that, if the zoom-lens would be gradually rescaled, the velocity of rescaling is not constant. On the other hand, changing the size of the zoom-lens to a greater extent might cost more time, but the change in size might have been too small in the present stimulus to influence RT significantly.

In sum, including conditions in which attention could be rescaled from and to the middle level led to complementary results to the previous study. Analyses of the RT suggested that rescaling to the middle level resulted in an overshoot in the extent of rescaling. The disappearance of this effect with prolonged stimulus duration could indicate that the accurateness of rescaling depends on the time available. In addition, the lack of RCCs for zooming-out from the small to large level, which contrasted with results of experiment 2, suggested that factors like strategy or practice can influence rescaling mechanisms.

Both the extent and the direction of rescaling appeared to influence rescaling behavior, although their influences did not interact. More errors were made when the zoom-lens was rescaled to a greater extent (i.e., to the opposite instead of an adjacent level). Furthermore, zooming-in led to more errors than zooming-out or, when stimulus duration was prolonged, to an increase in mean RT compared to zooming-out. In line with experiment 2, zooming-in was also associated with higher error RCCs. However, since differences in figure-ground segmentation difficulty between the levels of the used stimulus might have contributed to these diverging results for zooming-in and zooming-out, additional studies with other stimuli would be needed to confirm that the zoom-lens is indeed differently expanded than contracted.

Experiment 4: Scaling attention in the newly developed three-level compound stimulus is not affected by occlusion.

The results of experiment 1 suggested that attention could be accurately scaled to the middle and small level, but less accurately to the large level if the distance between the large and middle level was too small. To ensure that differences in the degree of occlusion between the levels did not affect results, the three-level compound stimulus in the present experiment was displayed in a dynamic fashion, which eliminated the occlusion between letters.

Methods

Participants

Twelve healthy volunteers (4 males; mean age: 21.7 years; standard error: 1.0 years) with normal or corrected-to-normal visual acuity were paid for participation in this study. All participants (except one) were right-handed and none of them took part in the previous experiments. They gave their informed consent to participate in the study, which was approved by the local Ethical Committee. One participant, having a mean RT that was more than two standard deviations above the group mean, was excluded from subsequent analyses.

Design

The design was identical to that of experiment 1, except that the three letters of the compound stimulus were presented in alternation instead of simultaneously. Each letter was presented for 16.6 ms in nine loops of large, middle and small letter presentation, creating a stable percept of the compound stimulus in which all letters were entirely visible. The presentation duration of this ‘dynamic’ version of the compound stimulus was 450 ms, which was (like all other characteristics) similar to the duration of the ‘static’ stimulus used in experiment 1. To ensure that the timing of the stimulus was correct, a (17”) CRT instead of LCD monitor was used and correct timing of stimulus presentation was checked. Data analysis was similar to that of experiment 1.

Results

Responses were overall slower (mean correct RT: 917 (± 32) ms) in this compared to the first experiment ($t(10) = 2.5$; $p = .032$), whereas accuracy (91.9(± 0.8)%) was similar in both experiments ($t(10) = 1.8$; $p = .109$). Similar to experiment 1, a significant interaction between stimulus size (i.e., the laM or smM stimulus containing the large or small middle level, respectively) and cued level was obtained from the RT analysis (Figure 7 and Table 4). Also the post-hoc tests per

stimulus size provided comparable results to the first experiment: When the laM stimulus was presented, reaction times were slower for attending the large ($t(10) = 4.7$; $p = .001$) or the middle ($t(10) = 7.1$; $p = .00004$) compared to the small level. In addition, attending the middle level was slower compared to attending the small level ($t(10) = 3.3$; $p = .009$), in the smM stimulus. In contrast to experiment 1, the RT difference between responses to the large level in the laM compared to the smM stimulus reached significance ($t(10) = 4.6$; $p = .001$).

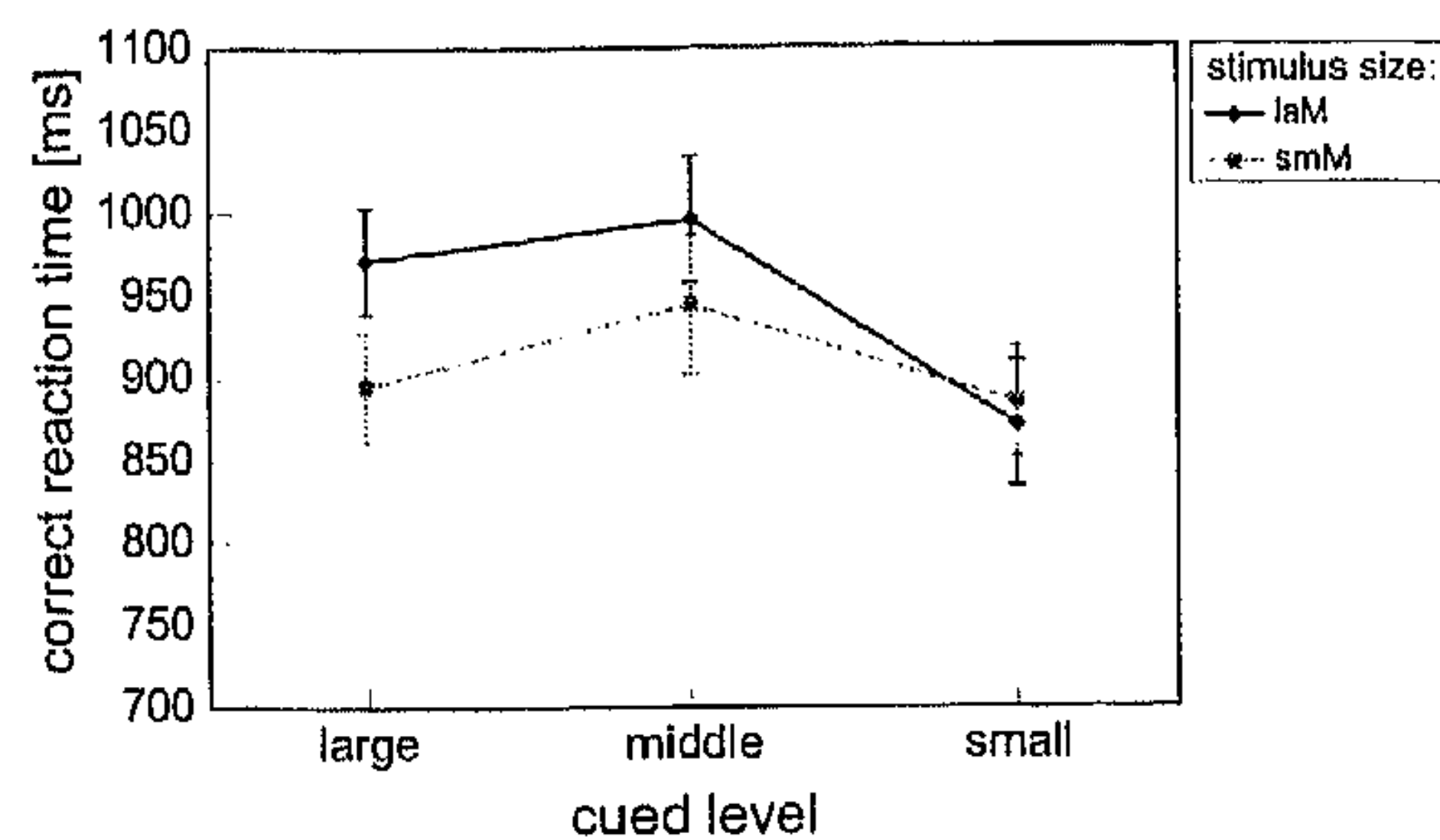


Figure 7. Mean reaction times of correct responses as a function of cued level and stimulus size. Error bars indicate SEM.

Furthermore, a main congruency effect on RT was found, which was caused by slower responses to stimuli with an incongruent middle ($t(10) = 2.9$; $p = .008$; one-sided) or incongruent small level ($t(10) = 4.5$; $p = .0006$; one-sided) compared to stimuli in which all levels were congruent.

Table 4. F-values, degrees of freedom (df), and Greenhouse-Geisser corrected p-values for the main and interaction effects of the mean correct reaction time (RT; left columns) and error percentage (Error; right columns) analyses. Abbreviations: Size: stimulus size; RC: response congruency. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.005$.

Factors	RT		Error	
	F (df)	p	F (df)	p
Size	16.3 (1,10)	0.002***	21.1 (1,10)	0.001***
Cued Level	13.9 (2,20)	0.001***	11.1 (2,20)	0.002***
RC	4.2 (3,30)	0.030*	16.1 (3,30)	0.000***
Size • Cued Level	6.6 (2,20)	0.010*	8.0 (2,20)	0.007**
Size • RC	0.5 (3,30)	-	2.8 (3,30)	-
Cued Level • RC	0.8 (6,60)	-	5.5 (6,60)	0.006**
Size • Cued Level • RC	0.9 (6,60)	-	1.1 (6,60)	-

Similar to experiment 1, all main effects in the error percentage analysis were significant (Table 4). Although (unlike experiment 1), the three-way interaction was not significant, two other interactions were significant whereas the third interaction showed a tendency towards significance ($p = .060$). To interpret the most important effects of these interactions (Figure 8), the RCC of each condition was computed and tested similar to experiment 1. When the laM stimulus was presented, significant RCCs were present when the large level was cued and the large (mean RCC = 17.6%; $t(10) = 4.8$; $p = .0005$; one-sided) or middle (mean RCC = 15.9%; $t(10) = 4.4$; $p = .0005$; one-sided) level was incongruent, whereas none of the other RCCs reached significance. The error RCCs that are solely obtained when the large level is attended in a laM stimulus with an incongruent middle and large level and not in any other case, is in line with the results of experiment 1.

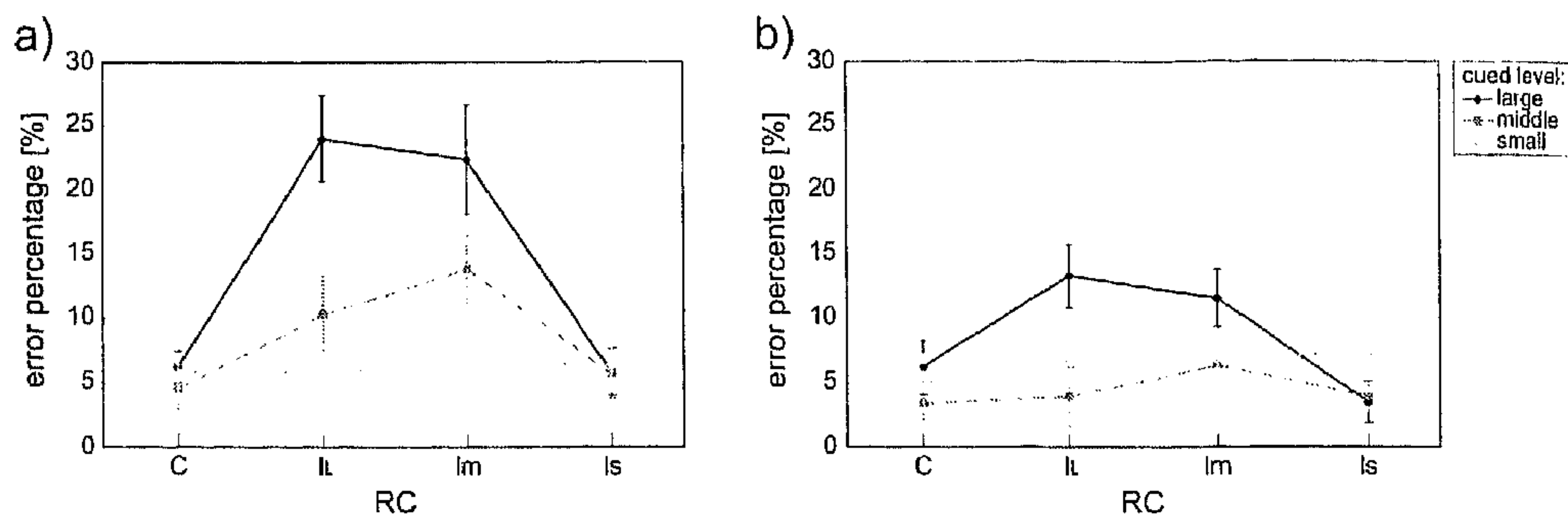


Figure 8. Mean error percentages as a function of response congruency level (C: all levels congruent, Il: incongruent large level, Im: incongruent middle level, Is: incongruent small level) and cued level separately shown for the a) laM stimulus and b) smM stimulus. Error bars indicate SEM.

Discussion

Overall, responses were slower in this experiment compared to experiment 1, in which an identical task was performed with a static version of the three-level compound stimulus. However, all other results matched the results of experiment 1 very well. Similar to experiment 1, letters at the small level were identified faster compared to the middle level in both stimulus sizes, whereas the mean RT for the small level was in addition faster than the large level in the laM but not in the smM stimulus. By replicating the findings of experiment 1, these results confirmed that the faster processing of stimuli at the small level was not caused by differences in occlusion between this level and the other levels. However, unlike experiment 1 (in which this difference did not reach significance), responses to the large level were significantly faster in the smM compared to the laM stimulus in the present experiment. In addition, responses were slower to stimuli with an incongruent middle or incongruent small level compared to stimuli in which all levels were congruent, irrespective of the cued level. These results suggested that scaling the zoom-lens solely to the large level was more difficult in the dynamic than the static version of the compound stimulus. Probably, although the percept of the three letters was stable in the dynamic stimulus, the flickering of the letters slowed responses. Likewise, stimulus regions in which the flickering was most salient (i.e., the regions at the small and middle level in which letters overlapped) might have captured attention, as suggested by RT RCCs of the middle and small level. Finally, results of the error analysis resembled those in experiment 1. In both experiments, attention could be precisely scaled to the small and middle level. Scaling to the large level could be quite precise, but only when the distance between the large and middle level was large enough. In sum, the strong convergence between the results of experiment 1 and 4 suggested that differences in occlusion between the levels do not affect the scaling of attention considerably.

General discussion & Conclusion

In a series of experiments the mechanisms of rescaling the attentional zoom-lens were explored using a newly developed three-level compound stimulus. Results of the first experiment revealed that attention could be accurately scaled to each of the three levels, given that the distance between the large and middle level was large enough. Only when attention was scaled to the large level of a stimulus with a relatively large middle level, interference between two levels was observed. These findings indicated that, in line with existing literature, attention can indeed be allocated in an annular fashion, although the width of the attentional annulus has a lower limit. The second study investigated whether the zoom-lens was continuously or discretely rescaled by presenting a zoom-symbol on the cued level that required participants to rescale attention to the opposite level (from large to small or v.v.) within the same compound stimulus. In stay trials, the small level interfered when the large level was attended. This suggested that although experiment 1 showed that attention can be accurately scaled to each level, participants are able to strategically divide their attention over two non-adjacent regions. However, when attention had to be rescaled from the large to the small level, the intermediate level strongly interfered. This indicated that the intermediate level cannot be “skipped” when attention is rescaled from the large to the small level, even though findings on stay trials suggested that some attention appears to be already deployed to the small level. Interference of the middle level was also observed when attention was zoomed-out from the small to the large level, although this interference was weaker than for zooming-in. Despite that the results of experiment 2 suggested that attention was gradually rescaled between the small and large level, no indications were found that rescaling attention to a larger extent (i.e., rescaling to the opposite level in the large compared to small stimulus) took more time. The third study explored the influence of the direction (i.e., zooming-in versus zooming-out) and the extent of rescaling (i.e., to the opposite level [between large and small] versus to the adjacent level [between large and middle or middle and small]) on rescaling mechanisms in more detail, by including conditions in which attention could be rescaled from and to the middle level. In line with experiment 2, zooming-in led to more errors and a higher error RCC compared to zooming-out. A higher error percentage for rescaling to the opposite compared to the adjacent level suggested that, not only the direction, but also the extent of rescaling influenced accuracy. Although the effects on accuracy were in the anticipated direction, these effects had been expected in the RT instead of the error analyses. Since results suggested that voluntary scaling of attention was quite time-consuming and the used stimulus duration might therefore have been too short to properly rescale attention, an additional experiment was performed with prolonged stimulus duration (but otherwise identical to experiment 3). In this experiment, both the direction as well as the extent of rescaling influenced RT, although the latter effect was very weak. Finally, a fourth study

was performed to control for occlusion confounds by using a 'dynamic' version of the compound stimulus in which none of the three levels was occluded by another level. The study replicated the main results of the first study, which suggested that the obtained results could not be explained by differences in occlusion between different levels.

Combining the results of the four experiments indicates that, although attention can be accurately scaled to each level in the used compound stimulus (experiment 1 and 4), the zoom-lens cannot "jump" from one level to the other level when attention is zoomed-in (experiment 2 and 3). Since the diverging findings for zooming-in and zooming-out are probably due to figure-ground segmentation differences between levels, visual context appears to influence rescaling of the zoom-lens. Moreover, when the large level was attended the small level interfered in (stay trials of) experiment 2 but not in experiment 1, which indicated that scaling mechanisms can be influenced by strategies. In addition, the optimization of zooming-out from the small to the large level that was observed in experiment 3, but not in experiment 2, suggested that also other factors like practice could influence rescaling behavior. This influence appeared to be that strong that zooming-out from the small to large level seemed to occur in a discrete fashion in experiment 3. Overall, these findings indicate that attentional scaling mechanisms are strongly influenced by both top-down factors, like strategy, as well as bottom-up factors, like visual context. This dynamic nature of rescaling mechanisms hampers general interpretations concerning behavior of the zoom-lens. Likewise, the nature of attentional shifting across the visual field is quite ambiguous given the discussions in literature (for a review, see Cave & Bichot, 1999). However, more recent studies suggested discrete instead of analog attention shifts (e.g., Chastain, 1992a, 1992b). In line with these findings, experiment 3 suggested that the zoom-lens can discretely be rescaled. However, since discrete rescaling was only observed for expanding the zoom-lens and, moreover, only when the stimulus always had the same size (experiment 3), discrete rescaling of the zoom-lens might only be feasible under optimal conditions. Future studies are necessary to explore the influence of the direction of rescaling, as well as the influence of top-down and bottom-up factors on rescaling in more detail. Similar to paradigms used to study attention shifts, the present paradigm does not allow the tracking of attention at every time point. The results of the present study suggested that the behavior of the zoom-lens is influenced by many factors and it might be that the fashion of rescaling can even differ during its time-course. For example, since the stimuli in the present experiments were always presented with spatial certainty, zooming-in from the large to the small level might have been strategically divided in two stages: in an initial stage, the annular zoom-lens encompassing the large level could have been discretely contracted/reshaped to an annular or circular zoom-lens covering the middle and small level to allow a coarse, attentive segregation of the small letter from its background, after which in a final stage, the zoom-lens would be contracted in smaller steps to encompass the

small level for accurate letter identification. Finally, one potential confound deserves mention. Similar to many other studies (e.g., Eriksen & Eriksen, 1974; LaBerge et al., 1991; McMains & Somers, 2005; Müller & Hübner, 2002), a letter identification task was used to investigate the behavior of spatial attention. Although the results are discussed in the context of rescaling spatial attention, object-based attention (e.g., Duncan, 1984; Vecera & Farah, 1994; for a review see Scholl, 2001) probably also played a role in the present experiments. Instead of a truly circular (or annular) shaped zoom-lens, it was therefore more likely that the actual applied attentional distribution was the result of an interaction between spatial attention (directed by the spatial cue) and object-based attention (e.g., Abrams & Law, 2000; Müller & Kleinschmidt, 2003), in which the borders of the zoom-lens were (coarsely) demarcated by the contours of the attended letter.

In sum, the present results indicate that attentional scaling mechanisms are dynamically influenced by both top-down factors, like strategy, as well as bottom-up factors, like visual context. Some results suggested that, under optimal circumstances, discrete scaling of the zoom-lens appears to be possible. However, both the direction as well as the extent of rescaling influenced scaling behavior of the zoom-lens. Future studies could investigate the top-down and bottom-up factors influencing scaling mechanisms in more detail, by for example varying the amount of training, the stimulus duration and the type of stimulus. In addition, the use of neurophysiological techniques would provide detailed information concerning the time-course of scaling the zoom-lens, whereas neural mechanisms of attentional scaling could be studied using functional magnetic resonance imaging.

Acknowledgments

We acknowledge Dirk Vorberg for developing the stimuli and Joel Reithler for useful discussions and comments on the manuscript. This research was supported by NWO Grant No. 402-01-632 to both authors.

References

- Abrams, R. A., & Law, M. B. (2000). Object-based visual attention with endogenous orienting. *Perception & Psychophysics*, 62(4), 818-833.
- Castiello, U., & Umiltà, C. (1990). Size of the attentional focus and efficiency components of processing. *Acta Psychologica*, 73, 195-209.
- Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: beyond a spotlight model. *Psychonomic Bulletin & Review*, 6(2), 204-223.
- Chastain, G. (1992a). Analog versus discrete shifts of attention across the visual field. *Psychological Research*, 54, 175-181.
- Chastain, G. (1992b). Time-course of sensitivity changes as attention shifts to an unpredictable location. *Journal of General Psychology*, 119, 105-111.

- Driver, J., Davis, G., Russell, C., Turratto, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition*, 80, 61-95.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 501-517.
- Egely, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, 123, 161-177.
- Egely, R., & Homa, D. (1984) Sensitization of the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 778-793.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 143-149.
- Eriksen, C. W., & Hoffman, J. E. (1974). Selective attention: Noise suppression or signal enhancement? *Bulletin of the Psychonomic Society*, 4, 587-589.
- Eriksen, C. W., & St. James, J. D. (1986) Visual attention within and around the field of focal attention: a zoom lens model. *Perception & Psychophysics*, 40(4), 225-240.
- Eriksen, C. W., & Murphy, T. D. (1987). Movement of attention focus across the visual field: a critical look at the evidence. *Perception & Psychophysics*, 42(3), 299-305.
- Greenwood, P. M., & Parasuraman, R. (1999). Scale of attentional focus in visual search. *Perception & Psychophysics*, 61(5), 837-859.
- Grice, G. R., Canham, L., & Boroughs, J. M. (1983). Forest before trees? It depends where you look. *Perception & Psychophysics*, 33(2), 121-128.
- Hoffman, J. E. (1980). Interaction between global and local levels of a form. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 222-234.
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, 43(3), 171-216.
- Juola, J. F., Bouwhuis, D. G., Cooper, E. E., & Warner, C. B. (1991). Control of attention around the fovea. *Journal of Experimental Psychology: Human Perception and Performance*, 17(1), 125-141.
- Kinchla, R. A., & Wolfe, J. M. (1979). The order of visual processing: "Top down", "bottom-up", or "middle-out". *Perception & Psychophysics*, 25(3), 225-231.
- LaBerge, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 371-379.
- LaBerge, D., Brown, V., Carter, M., Bash, D., & Hartley, A. (1991). Reducing the effects of adjacent distractors by narrowing attention. *Journal of Experimental Psychology: Human Perception and Performance*, 17(1), 65-76.
- Lamb, M. R., & Robertson, L. C. (1988). The processing of hierarchical stimuli: Effects of retinal locus, locational uncertainty, and stimulus identity. *Perception & Psychophysics*, 44, 172-181.
- Luo, Y.-J., Greenwood, P. M., & Parasuraman, R. (2001). Dynamics of the spatial scale of visual attention revealed by brain event-related potentials. *Cognitive Brain Research*, 12, 371-381.
- Martin, M. (1979). Local and global processing: The role of sparsity. *Memory and Cognition*, 7, 476-484.
- McMains, S. A., & Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron*, 42(4), 677-686.
- McMains, S. A., & Somers, D. C. (2005). Processing efficiency of divided spatial attention mechanisms in human visual cortex. *The Journal of Neuroscience*, 25(41), 9444-9448.
- Müller, N. G., Bartelt, O. A., Donner, T. H., Villringer, A., & Brandt, S. A. (2003). A physiological correlate of the "zoom lens" of visual attention. *The Journal of Neuroscience*, 23(9), 3561 - 3565.
- Müller, M. M., & Hübner, R. (2002). Can the attentional spotlight be shaped like a doughnut? Evidence from steady state visual evoked potentials. *Psychological Science*, 13, 119 -124.

- Müller, N. G., & Kleinschmidt, A. (2003). Dynamic Interaction of Object- and Space-Based Attention in Retinotopic Visual Areas. *The Journal of Neuroscience*, 23(30), 9812–9816.
- Murphy, T. D., & Eriksen, C. W. (1987). Temporal changes in the distribution of attention in the visual field in response to precues. *Perception & Psychophysics*, 42(6), 576-586.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353-383.
- Nebel, K., Wiese, H., Stude, P., De Greiff, A., Diener, H. C., & Keidel, M. (2005). On the neural basis of focused and divided attention. *Cognitive Brain Research*, 25(3), 760-776.
- Posner, M. L. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32A, 3-25.
- Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition*, 80, 1–46.
- Shulman, G. L., Remington, R. W., & MacLean, J. P. (1979). Moving attention through physical space. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 522-526.
- Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review*, 102(3), 503-532.
- Stoffer, T. H. (1993). The time course of attentional zooming: a comparison of voluntary and involuntary allocation of attention to the levels of compound stimuli. *Psychological Research*, 56, 14-25.
- Tsal, Y. (1983). Movements of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 523-530.
- Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, 123, 146-160.
- Vecera, S. P., Fievaris, A. V. & Filapck, J. C. (2004). Exogenous spatial attention influences figure-ground assignment. *Psychological Science*, 15(1), 20-26.
- Zenger, B., Braun, J., & Koch, C. (2000). Attentional effects on contrast detection in the presence of surround masks. *Vision Research*, 40, 3717–3724.

Chapter 5

The neural mechanisms of scaling the attentional “zoom-lens”

Based on: Peters, J. C. & Goebel, R. The neural mechanisms of scaling the
attentional “zoom-lens”. (*in preparation*)

Abstract

Although it has been shown that humans can modify the size of their spatial attention window, the neural mechanisms involved in the scaling of this attentional "zoom-lens" have not been established. The present rapid event-related functional magnetic resonance imaging study investigated the neural mechanisms involved in expanding and contracting the zoom-lens. Participants were cued to either attend the large or small letter of a three-level compound stimulus. In one-third of the trials, a symbol at the cued level indicated that attention had to be rescaled to the opposite level within the same compound stimulus (i.e., attention had to be "zoomed-in" from large to small or "zoomed-out" from the small to the large level). Rescaling the zoom-lens recruited a fronto-parietal network, which largely overlapped with the network known to be involved in shifting attention from one location to another. Zooming-in and zooming-out of attention led to highly similar activations, although the extent as well as the strength of these activations was larger for zooming-in. To study whether the intermediate level would be temporarily encompassed in the zoom-lens when attention was rescaled between the large and small level, response congruency between the different levels was varied. Activity in the dorsal anterior cingulate cortex was enhanced when attention was zoomed-out in a stimulus with a response incongruent intermediate and target level, but not when attention was zoomed-in. This indicated that the zoom-lens temporarily encompassed the middle letter during zooming-out, whereas the zoom-lens appeared to be discretely zoomed-in from the large to the small level. Apparently, the fashion of rescaling differs between zooming-in and zooming-out, although both are mediated by similar fronto-parietal networks.

Introduction

Directing visual attention to a region in the visual field, in which the presentation of a target is expected, facilitates processing of the subsequently presented target (e.g., Eriksen & Hoffman, 1974; Posner, 1980). Besides this ability to shift the focus of attention through the visual field, the zoom-lens theory of attention (Eriksen & St. James, 1986) suggests that the size of the focus can be adjusted to match the spatial scale of a relevant region. Empirical results showed that the size of the attentional focus ("zoom-lens") can indeed be constricted or expanded according to task demands, whereas the size of the zoom-lens is negatively correlated with processing efficiency (Castiello & Umiltà, 1990; Eriksen & St. James, 1986; LaBerge, 1983; Luo, Greenwood, & Parasuraman, 2001; McMains & Somers, 2005; Müller, Bartelt, Donner, Villringer, & Brandt, 2003a). In addition to shifting the attentional focus, scaling the focus also appears to play an important role in visual search (Greenwood & Parasuraman, 1999) and target detection (Hopf et al., 2006). However, compared to the vast amount of psychophysical studies investigating covert attention shifts, relatively few experiments have been conducted on the mechanisms underlying the scaling of the attentional focus. For example, several studies explored whether the attentional focus shifts from one location to another in an analog (i.e., covering intermediate locations along the way) or discrete (i.e., 'jumping' from the first to the second location) manner (e.g., Chastain, 1992a, 1992b; Murphy & Eriksen, 1987; Shulman, Remington, & MacLean, 1979; Sperling & Weichselgartner, 1995; Tsal, 1983). However, to our knowledge it has not been studied yet whether the zoom-lens is continuously or discretely scaled. Likewise, although the neural mechanisms controlling attention shifts have been well-studied (e.g., Beauchamp, Petit, Ellmore, Ingelholm, & Haxby, 2001; Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000; Gitelman et al., 1999; Hopfinger, Buonocore, & Magnus, 2000; Yantis et al., 2002), much less is known concerning the neural mechanisms involved in scaling the zoom-lens. Although neural reflections of the zoom-lens have been observed in the visual cortex (Luo et al., 2001; McMains & Somers, 2005; Müller et al., 2003a), the neural mechanism involved in controlling the scaling of the zoom-lens are largely unknown. We are only aware of a single neuroimaging study (Müller et al., 2003b) to date addressed activations in frontal and parietal areas when attention was covertly oriented to a cued area that varied in size over trials. However, in this study the zoom-lens did not only have to be scaled during cue presentation, but first had to be shifted towards the cued area. In addition, after shifting and scaling attention, the established attentional setting had to be maintained throughout a relatively long cue presentation period (of 4, 7, or 10 s). In the present study, a paradigm was used which was more strongly directed to the attentional rescaling process per se. As illustrated in Figure 1, participants were cued to either attend to the large or small letter of a newly developed three-level compound stimulus (Vorberg, 1992; unpublished; see Chapter 4 for details).

One-third of the trials were “zoom” trials, in which a letter-like symbol at the cued level indicated that attention had to be redirected to the opposite level (from large to small or vice versa) within the same compound stimulus. By contrasting these trials with other trials in which attention could stay focused on either the large or the small level, the rescaling of the zoom-lens could be studied separately from other processes like shifting attention (since the cued regions were concentric, the attentional focus did not have to be moved during rescaling) and maintaining the established zoom-lens setting (since immediately after rescaling, the target letter could be processed).

In addition to investigating the neural mechanisms involved in rescaling the zoom-lens, the use of a stimulus containing three levels enabled us to explore whether the intermediate level can be “skipped” when attention is rescaled from the large to the small level or vice versa. If the zoom-lens could not be discretely rescaled between the large and small level, the zoom-lens would temporarily encompass the intermediate level, whereas no attention would be deployed to the intermediate level if the zoom-lens could “jump” between the small and large level. The allocation of attention to the intermediate level was investigated by systematically varying response congruency between the intermediate level and the level to which attention had to be rescaled (further referred to as the target level). Based on findings in flanker tasks (Eriksen & Eriksen, 1974), we hypothesized that, when the letters at the intermediate and target level were coupled to a different instead of the same response, response congruency costs would be present if the zoom-lens were continuously rescaled, whereas response congruency costs should be (close to) absent if the zoom-lens were discretely rescaled. Therefore, we carefully examined the potential influences of response conflict on neural responses in the anterior part of the rostral cingulate zone (as defined by Picard & Strick, 2001; this area will be further referred to as dorsal anterior cingulate cortex or dorsal ACC), which is known to be involved in response conflict monitoring (e.g., Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Picard & Strick, 2001; Ridderinkhof, Ullsperger, Crone, Nieuwenhuis, 2004). The present study was based on a series of psychophysical experiments (described in Chapter 4), in which the same type of paradigm was used. Importantly, these experiments indicated that attention could be accurately scaled to each level separately (i.e., without allocating attention to the other levels) and that scaling was not influenced by the presence of occlusion in the stimulus. Moreover, in one of these experiments, which used an almost identical paradigm (experiment 2 in Chapter 4), response congruency costs were observed when attention was rescaled from the large to the small level or the other way around, suggesting that the zoom-lens could not “jump” between the large and small level. Note that each participant of the present study took part in this behavioral experiment as training for the fMRI session. In this training session, eye movements were monitored to ensure that rescaling of the zoom-lens was not accompanied by overt shifts of attention.

Methods

Participants

Ten healthy volunteers (4 males; mean age 27.4 years; standard error: 2.1 years) with normal or corrected-to-normal visual acuity were paid for participation in this study. All participants were right-handed and gave their informed consent to participate in the study, which was approved by the local Ethical Committee. Due to file corruption, eye movement data of three subjects were lost.

Design

As illustrated in Figure 1, each trial began with the presentation (500 ms) of a large ($7.0^\circ \times 3.0^\circ$) or small ($1.75^\circ \times 0.75^\circ$) outlined rectangle. The size of this rectangle matched the size of either the large or the small character of the three-level compound stimulus that was subsequently presented (300 ms) after a variable cue-delay (3 ± 1.5 s). That is, the size of the rectangle indicated which level of the stimulus should be attended and participants could use this cue to optimize their spatial attention window for the upcoming stimulus. The small and large level were cued equally often. The compound stimulus consisted of a superimposed large ($7.0^\circ \times 3.0^\circ$), medium ($3.5^\circ \times 1.5^\circ$), and small ($1.75^\circ \times 0.75^\circ$) character. Each character could be the letter B, N, H, or S, or a so-called ‘zoom symbol’. Examples of this letter-like zoom symbol presented at the large or small level of the compound stimulus are displayed on the right of Figure 1 in the middle and lower stimulus, respectively. Two types of trials were included, as illustrated in the right part of Figure 1. In so-called *stay*-trials, a letter was presented at the cued level of the compound stimulus and participants were required to report the identity of this letter with a button-press. Half of the participants responded to the letters B and N with the right index finger and to the letters H and S with the right middle finger, while this stimulus-response coupling was reversed for the other half of the participants. In contrast to *stay*-trials, in *zoom*-trials the zoom-symbol was presented at the cued level, requiring participants to redirect their attention to the opposite level (from large to small or vice versa) and report the identity of the letter presented at that level. Each of these two trial types contained two subtypes (since the large or small level could be cued) resulting in four conditions, labeled *stay-large* and *stay-small* when attention did not have to be rescaled to another level and *zoom-in* (from the large to the small level) and *zoom-out* (from the small to the large level) if attention had to be rescaled.

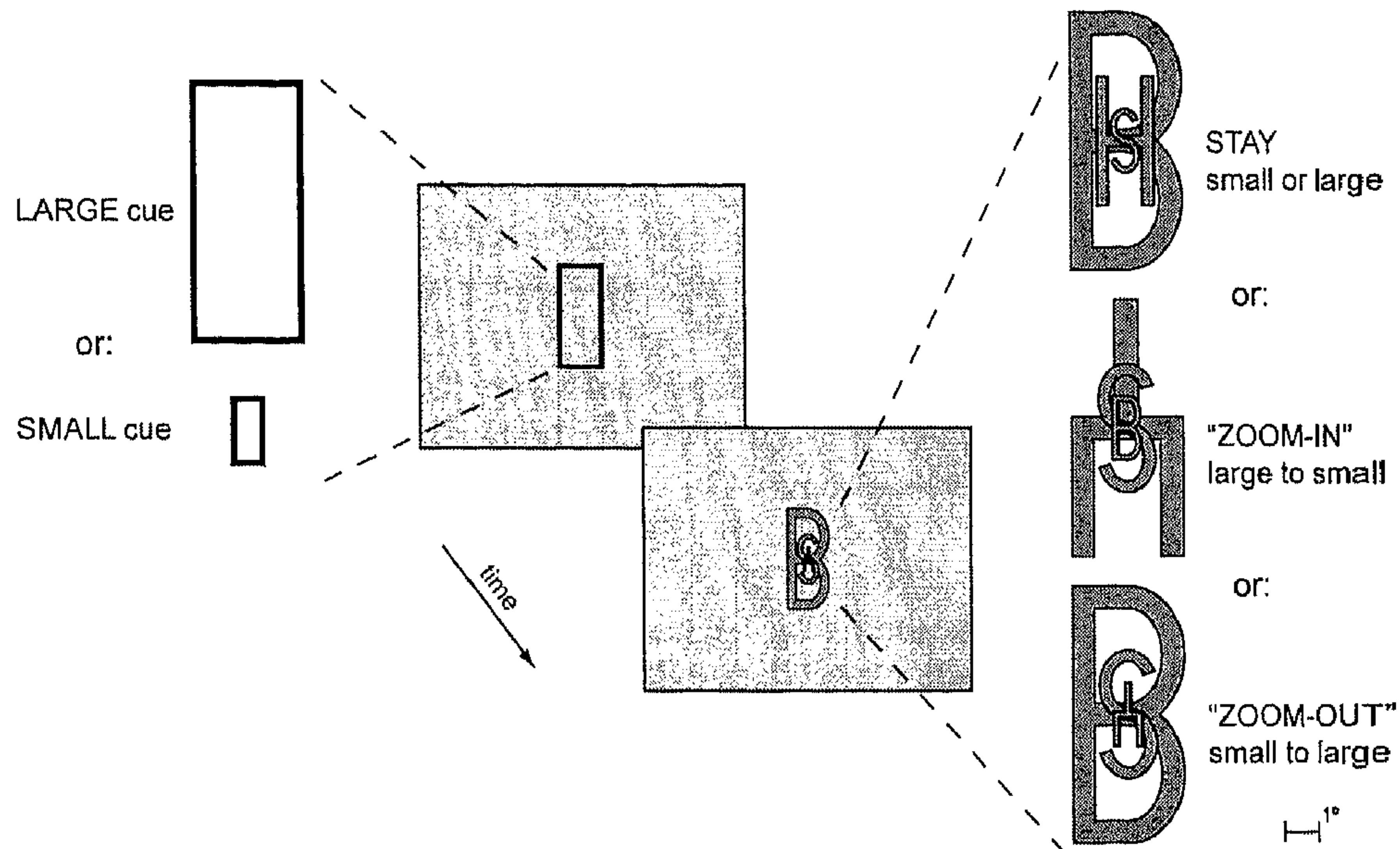


Figure 1. Schematic illustration of trial design. Each trial started with the presentation of a large or small outlined rectangle indicating the level that should be attended in the three-level compound stimulus that was subsequently presented after a variable cue-delay. Four types of trials were included, illustrated on the right (stimuli are drawn to scale). If a letter was present at the cued level of the compound stimulus, the identity of this letter should be reported (stay condition). In contrast, if a “zoom” symbol was presented at the cued level (depicted on the large level of the upper compound stimulus and the small level of the middle compound stimulus) attention had to be redirected to the opposite level and the identity of the letter at this target level should be reported (zoom-in/zoom-out conditions). See text for further details.

Response congruency between the three levels was varied over trials. In stay-trials, the letters on the three levels could all be coupled to the same response (congruent condition), or one letter could be coupled to an alternative response. The latter case constituted three types of incongruent conditions, according to the level at which the incongruent letter was presented (i.e., incongruent small level, incongruent middle level, or incongruent large level). Note that the labeling of the response congruency conditions is independent of the level that needs to be attended. For example in the ‘incongruent large level’ condition, the large letter of the compound stimulus is always coupled to a different response than the middle and small letter. Therefore, the large letter is considered as the incongruent item, irrespective of the attended level. In zoom-trials, the compound stimulus contained only two letters, in addition to the neutral (i.e., not coupled to a response) zoom symbol at the cued level. These two letters on the middle and target level could be coupled to the same response (congruent condition) or to different responses (incongruent condition). After presentation of the compound stimulus a fixation cross was presented for a variable duration (intertarget interval: 5.7 ± 1.5 s). All stimuli were presented in the middle of the screen on a grey background. The zoom symbol was never presented at an uncued level. Conditions were first-order

counterbalanced. In addition, the combination of the type of cue and the corresponding cue-delay length as well as the combination of the trial type and the intertarget interval length were first-order counterbalanced to optimize temporal jitter. Index and middle finger responses were randomized but occurred equally often, whereas presented letters were randomly chosen from the set that was associated with the selected responses. Throughout the experiment, participants were reminded to fixate the fixation cross, especially during stimulus presentation, and to respond as fast and accurately as possible.

All participants took part in a behavioral experiment prior to the scanning session. The task in this experiment was identical to the task used in the fMRI experiment (the fMRI session task), except that next to the compound stimulus with a similar size as in the fMRI session task, an additional compound stimulus was introduced that was twice as large. In contrast to the fMRI session task, participants received feedback concerning the correctness of their response after every trial. Finally, although the cue-delay (500 ± 200 ms) and intertrial interval (2200 ± 200 ms) were jittered, similar to the fMRI session task, the average duration of both was shortened in order to increase the number of trials per condition. The behavioral experiment consisted of three phases: First, stimulus-response coupling was trained. Note that each participant used the same stimulus-response association in the behavioral and the fMRI session task. Then, participants were familiarized with the task, first by using a version of the task in which every phase of the trial lasted twice as long compared to the main task, followed by training the task with a trial timing identical to the main task. Finally, each participant performed four runs of the main task (576 trials) in which eye movements were monitored using an Eyelink tracker with a sampling rate of 250 Hz (SensoMotoric Instruments GmbH, Teltow, Germany). Eye movements were categorized as saccades when the movement velocity exceeded $35^\circ/\text{s}$ or when the movement acceleration was higher than $9500^\circ/\text{s}^2$. The behavioral experiment was part of a series of psychophysical experiments discussed in Chapter 4 of this thesis. Therefore, only results relevant for the fMRI study will be discussed in this Chapter. Accordingly, all mentioned effects resulted from analyses that included only the participants that took part in the present study. Moreover, only the trials in which a compound stimulus with the same size as in the fMRI session task was presented, were included in the analyses.

Image acquisition

Echo-planar images (T2*-weighted; 64x64 imaging matrix, 25 slices, voxel size: $3.5 \times 3.5 \times 3.5$ mm³, no gap, TR/TE = 1.5 s/30 ms, FA = 75°) covering almost the whole brain (except cerebellum) were collected on a 3-T Siemens Magnetom Allegra head-scanner (Siemens Medical Systems, Erlangen, Germany) using a standard head coil. Functional data were aligned to a T1-weighted high-resolution anatomical image (Magnetization-Prepared Rapid Acquisition Gradient Echo

(MPRAGE) sequence; TR/TE = 2.3 s/3.93 ms, FA = 8°). Participants viewed the stimuli, projected onto a frosted screen using a liquid crystal display projector (PLC-XT11-16, Sanyo North America Corporation, San Diego, USA), via a mirror mounted to the head coil. Stimuli were presented and responses were recorded using the Presentation software package (version 9.20; Neurobehavioral Systems, San Francisco, USA). Stimulus presentation was synchronized with MR data acquisition. Each participant performed 288 trials in total in eight runs. Immediately prior to the fMRI measurement, participants were trained in the task. In addition, each participant took part in a separate training session, in which eye movements were monitored.

Data analysis

Behavioral data

Only correct responses that occurred within the response time window defined as the interval from three standard deviations below to three standard deviations above the mean reaction time of each condition of each participant were included in subsequent analyses. Stay and zoom-trials were independently analyzed, since they differed in the number of congruency levels. Reaction times (of correct responses) and error rates of stay-trials were submitted to separate two-way repeated measures analyses of variance (ANOVAs) with attention condition (stay-large, stay-small) and response congruency (all levels congruent, incongruent small level, incongruent medium level, or incongruent large level) as factors. Correct reaction times and error rates of zoom-trials were entered into similar ANOVAs with attention condition (zoom-in, zoom-out) and response congruency (two congruent or two incongruent levels) as factors.

fMRI data

Preprocessing of the individual datasets included slice scan time correction, linear trend removal, high-pass filtering, three-dimensional motion correction, transformation into Talairach space (Talairach & Tournoux, 1988), and cortex reconstruction, inflation and flattening as implemented in the BrainVoyager QX software package (version 1.7.9; Brain Innovation, Maastricht, the Netherlands). The first two volumes of each run were discarded to remove T1 saturation effects. No spatial smoothing was applied to the functional data, which were interpolated to a 3 x 3 x 3 mm³ voxel target resolution.

Three types of analyses were performed to study the rescaling mechanisms of the zoom-lens. First, a whole brain voxel-wise multiple regression analysis was performed to establish the network involved in rescaling the zoom-lens. The design matrix included four predictors for the different attentional conditions (stay-large, stay-small, zoom-in, zoom-out) during compound stimulus presentation. In addition, two predictors were included to conjointly represent the cue and cue-delay of the large and small cue. Finally, incorrect trials were

separately modeled with a seventh predictor. Each predictor's boxcar function was convolved with a theoretical Two Gamma hemodynamic response function. The second analysis focused on potential differences between expanding and contracting the zoom-lens within the rescaling network obtained in the first analysis. A multi-subject deconvolution analysis was carried out to estimate neural responses to the compound stimulus, with minimal contamination of neural responses to temporally adjacent events (Serences, 2004; Glover, 1999). The deconvolution design matrix represented the same conditions as the first design matrix, except that the cue/cue-delay of the large and small cue were jointly modeled to reduce the number of predictors that needed to be estimated. Each condition was incorporated in the design matrix with six delta-function predictors, modeling the first part (0-7.5 s) of the hemodynamic response for each of the six time points independently. To quantify potentially diverging responses to zoom-in and zoom-out trials in the rescaling network, mean beta weights corresponding to the peak of the estimated hemodynamic responses were compared. Finally, a second multi-subject deconvolution analysis was performed to investigate whether rescaling occurred in a continuous or discrete fashion. The deconvolution design matrix represented six conditions. Four conditions modeled the incongruent and congruent versions of the compound stimulus, separately for the zoom-in and zoom-out condition. Stay and error trials were jointly modeled with a single predictor, since both conditions were not of interest and joining them considerably reduced the number of predictors that needed to be estimated. For the same reason, the large and small cues were modeled with a single predictor. Similar to the first deconvolution design matrix, each condition was incorporated in the design matrix by including six delta-function predictors. To study the influence of response congruency on rescaling the zoom-lens in detail, deconvolution analysis was confined to the dorsal ACC, which is known to be involved in conflict monitoring (e.g., Botvinick et al., 1999). Volumes of Interest (VOIs) in the dorsal ACC were identified for each participant individually by contrasting all predictors versus baseline (using the same design matrix as in the whole brain analysis). From the resulting statistical map, activated areas in the ACC were selected using a strict threshold (on average $p(\text{Bonferroni corrected}) < 0.0005$, but see Table 2 for subject-specific details). Finally, mean beta weights, corresponding to the peak of the estimated hemodynamic VOI responses in incongruent and congruent zooming trials, were contrasted to quantify potential activation differences.

Results

Behavioral results

As shown in Figure 2, participants were slower ($t(9) = 10.8$; $p = .000002$; Figure 2a) and made more errors ($t(9) = 4.2$; $p = .002$; Figure 2b) when attention had to be rescaled.

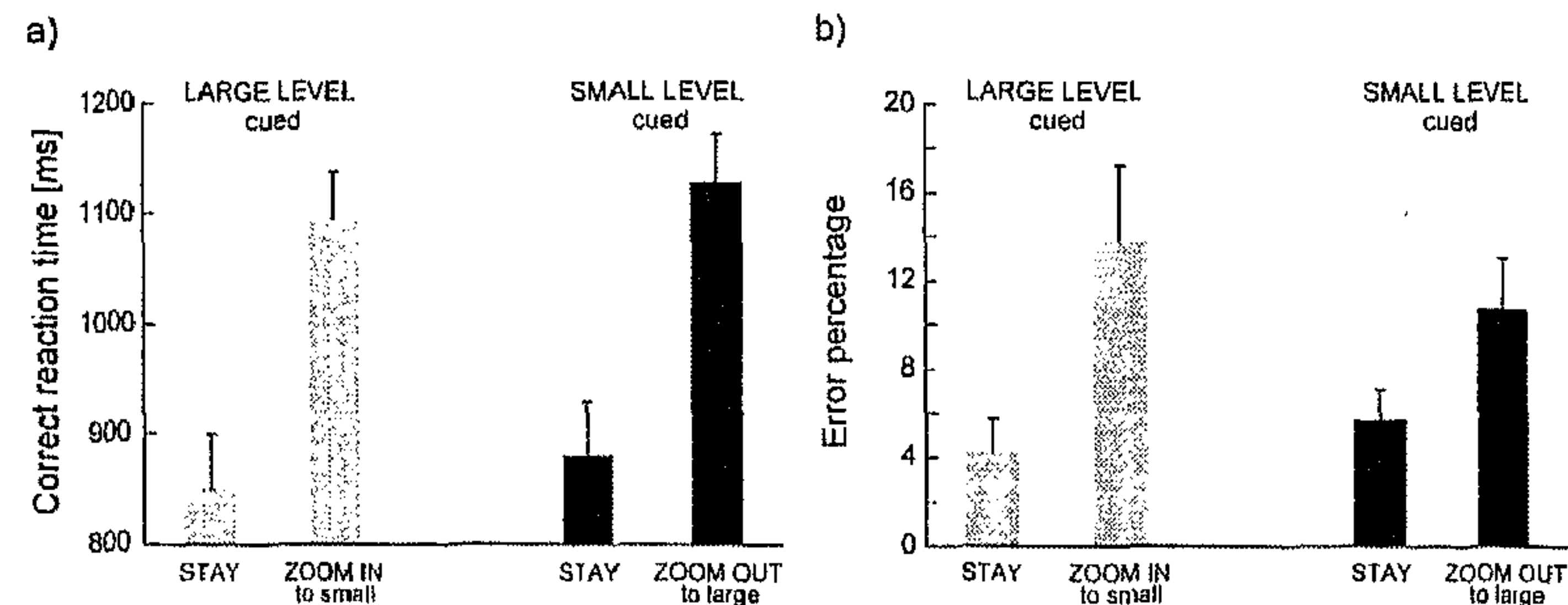


Figure 2. Mean reaction times (a) and error percentages (b) for the different attentional conditions. Rescaling of the zoom-lens leads to slower and less accurate responses. Error bars indicate standard errors of the mean (SEM).

In addition, participants were slower ($F(1,9) = 7.5$; $p = .023$) but equally accurate ($F(1,9) = 3.1$; $p = .111$) when attending to the small compared to the large level of the stimulus in stay-trials. No other effects were obtained. Although zoom-trials did not show a significant effect of response congruency on reaction time (RT; $p > .15$) or error percentage ($p > .05$), the response congruency cost was computed to compare behavioral and neural indices of response congruency. The RT response incongruency cost (Figure 3a) was computed for each participant by subtracting mean RT in the congruent condition from the mean RT in the incongruent condition for the zoom-in and the zoom-out condition separately. The error response congruency cost was computed in a similar fashion (Figure 3b).

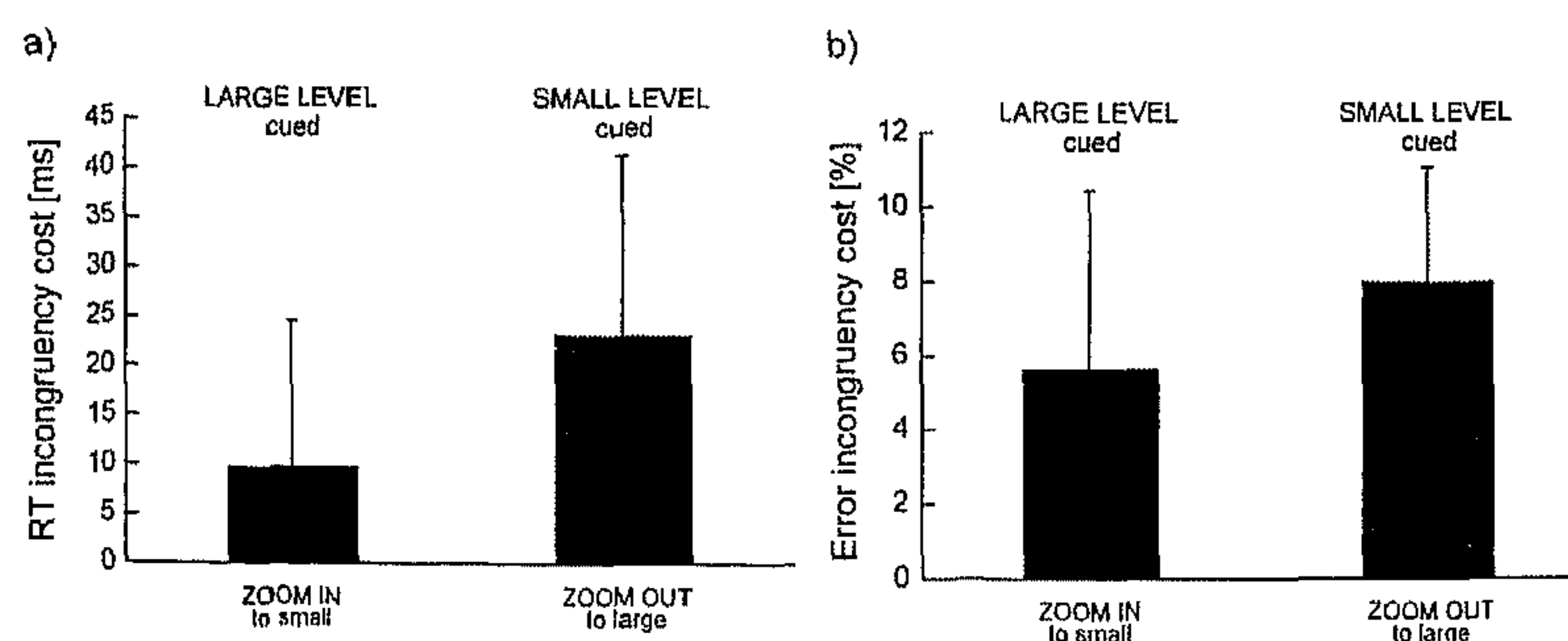


Figure 3. Response incongruency costs for mean reaction times (RT; a) and error percentages (b) for the different zoom conditions. Error bars indicate SEM. See text for further details.

As illustrated in Figure 3, there appeared to be a response congruency cost in both the zoom-in (mean (\pm standard error); RT cost: 10 (± 15) ms; error cost: 5.7 (± 4.8)%) as well as the zoom-out (RT cost: 23 (± 18) ms; error cost: 8.0 (± 3.1)%) condition. However, due to large inter-subject variability, (one-sided) planned comparisons between the incongruent and congruent trials of the zoom-in and the zoom-out condition revealed only a significant difference for accuracy in the zoom-out condition ($t(9) = 2.6$; $p = .015$). In contrast, the lower accuracy in incongruent compared to congruent trials did reach significance in both the zoom-

out (difference = 4.2%; $t(9) = 2.0$; $p = .037$) as well as the zoom-in (difference = 7.1%; $t(9) = 3.3$; $p = .005$) condition in the behavioral training session. Additionally, analyses of eye movement in the behavioral training session did not reveal differences in eye-movements between rescaling and static zoom-lens conditions. Neither the amount of saccades ($t(6) = 0.3$; $p = .759$) nor the overall deviation from fixation cross ($t(6) = 0.2$; $p = .842$) differed between stay and zoom-trials.

fMRI results

Rescaling the zoom-lens

When a zoom symbol was presented on the cued level, participants were required to rescale their attention to the opposite level. Expanding (zooming-out) or contracting (zooming-in) the size of the zoom-lens recruited a bilateral network of frontal and parietal areas, revealed by contrasting zoom and stay trials ($p(\text{Bonf}) < 0.02$; see Figure 4).

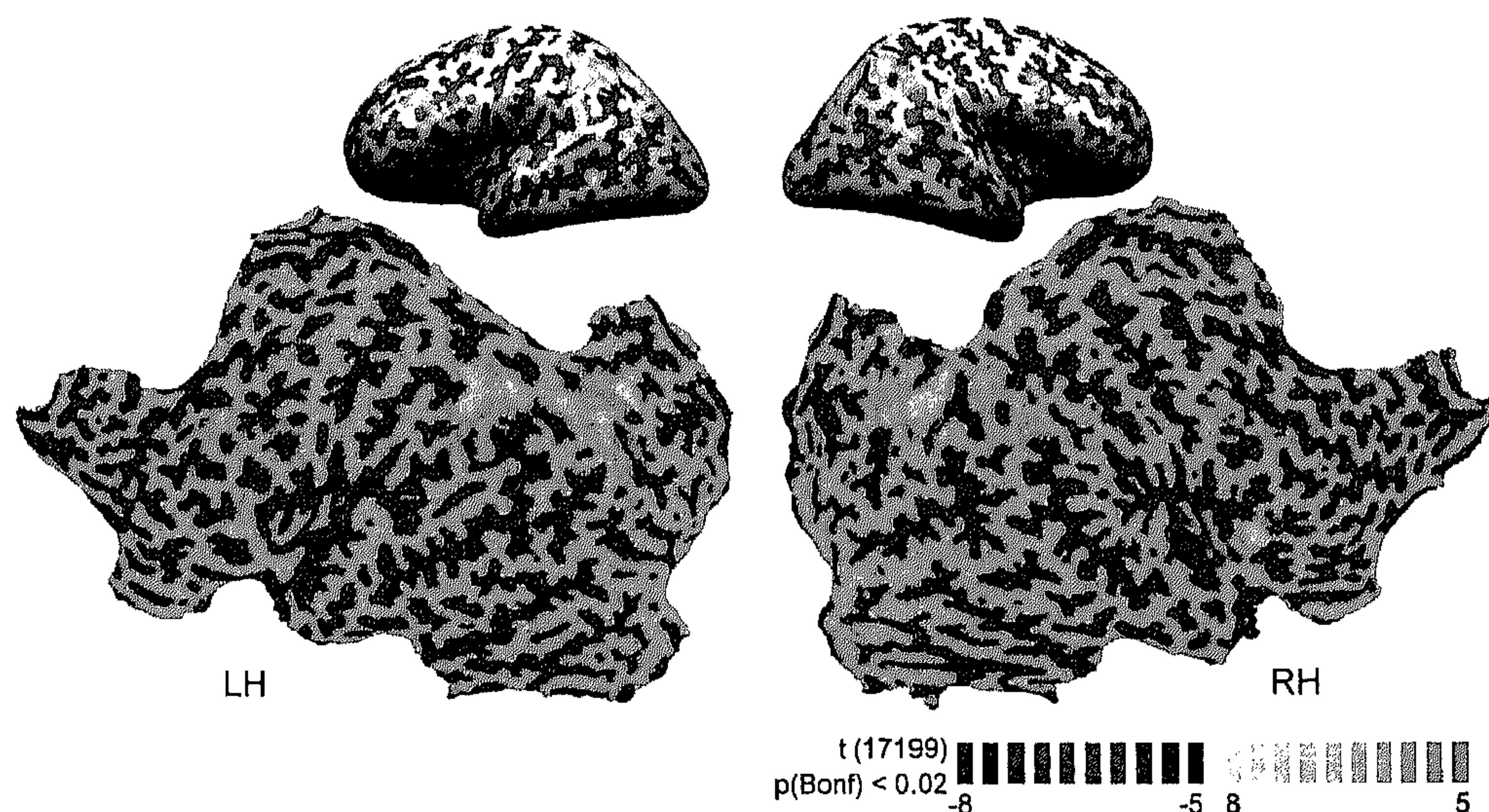


Figure 4. Rescaling the zoom-lens recruited a fronto-parietal network, which is projected on the inflated (upper figures) and flattened (lower figures) representation of a participant's grey-white matter boundary surface (LH: left hemisphere; RH: right hemisphere; dark grey: concave; light grey: convex cortical surface folding). Colored areas show a stronger activation when attention had to be rescaled to a different level compared to keeping attention focused on the same level, as resulting from a conjunction analysis in which zoom-in and zoom-out trials were contrasted with stay-large and stay-small trials with the constraint that the activations elicited by the included conditions had to be stronger than activations in the baseline condition ($p(\text{Bonf}) < 0.02$; only activation patches of more than 25 mm^2 are shown). Bonf: Bonferroni corrected.

Frontal activations related to rescaling the zoom-lens were found among others in the inferior frontal junction (IFJ; i.e., the junction of the inferior frontal sulcus

(IFS) and the inferior precentral sulcus [PrCS]), the left⁵ junction of the superior frontal sulcus (SFS) and the superior PrCS, which corresponds to the putative location of the frontal eye fields (FEF; e.g., Blanke et al., 2000), the dorsal ACC, the middle part of the middle frontal gyrus (mMFG) located in the dorsolateral prefrontal cortex (DLPFC), and the anterior insula (aINS); Table 1a). In addition, a large part of the parietal cortex was activated in both hemispheres. The left and right intraparietal sulcus (IPS) showed enhanced activity in the anterior (aIPS) as well as in the posterior part (pIPS). From the pIPS, branches of activation extended into the superior parietal lobe (SPL), the precuneus (PrCu) and the transverse occipital sulcus (TOS), especially in the left hemisphere.

a)

contrast	zoom > stay							
area	left hemisphere				right hemisphere			
	x	y	z	size	x	y	z	size
IFS/PrCS (IFJ)	-40	1	31	264	41	2	33	1638
mMFG	-45	21	36	1758	45	21	33	491
SFS/PrCS (FEF)	-29	-5	55	804	-	-	-	-
ACC	*	*	*	*	1	16	45	1992
aINS	-31	19	8	937	32	21	7	1141
aIPS	-36	-48	42	4572	38	-45	40	2150
pIPS/TOS/SPL/PrCu	-16	-69	40	3903	22	-67	40	4182

b)

contrast	zoom in > stay large								zoom out > stay small							
area	left hemisphere				right hemisphere				left hemisphere				right hemisphere			
	x	y	z	size	x	y	z	size	x	y	z	size	x	y	z	size
IFS/PrCS (IFJ)	-41	3	32	2344	43	3	32	4160	-	-	-	-	43	4	34	810
mMFG	-44	22	35	3585	43	24	29	1564	-46	20	37	1106	46	14	35	869
SFS/PrCS (FEF)	-28	-5	56	2344	-	-	-	-	-30	-5	56	524	-	-	-	-
ACC	*	*	*	*	0	15	43	5581	*	*	*	*	2	16	48	1124
aINS	-35	17	9	2709	35	20	7	2570	-30	19	8	672	32	21	7	880
aIPS	-37	-47	42	7755	38	-45	41	5860	-41	-46	45	2256	43	-46	42	857
pIPS/TOS/SPL/PrCu	-19	-67	41	10115	21	-67	40	9115	-20	-65	42	4204	22	-67	39	2972

Table 1. Talairach coordinates (in mm) of the centre of gravity and the size (in mm³) of clusters activated in the group analysis, identified by a) contrasting the zoom and stay condition conjointly (rescaling network; $p(\text{Bonf}) < 0.02$) or by b) contrasting the zoom-in and the stay-large (zoom-in network; $q(\text{FDR}) < 0.01$) and contrasting the zoom-out and stay-small (zoom-out network; $q(\text{FDR}) < 0.01$) conditions separately. To minimize the possibility of including clusters of artificial activations, all contrasts were analyzed in conjunction with a second contrast of all activations against baseline. Both tables only show areas of which the left or the right cluster has a size of at least 300 mm³. In addition, Table 1b shows only activation clusters revealed by both contrasts. However, only a few small activation clusters were found that were unique for the zooming-in or zooming-out network. Note that the volumes of the activation clusters are larger in the zoom-in compared to the zoom-out network. See text for abbreviations. *Although activation in the ACC was lateralized to the right hemisphere, a part of the left hemisphere was also included in this cluster. Bonf: Bonferroni corrected. FDR: False-Discovery-Rate.

To explore potential influences of the direction of rescaling (i.e., zooming-in or zooming-out), the networks recruited by zooming-in and by zooming-out were

⁵ A right hemispheric activation symmetrical to the left FEF activation was obtained when the statistical threshold was decreased to $q(\text{FDR}) < 0.001$.

also analyzed separately, by contrasting zoom-in with the stay-large and zoom-out with the stay-small condition (with the constraint that the activations elicited by the included conditions had to be higher than baseline). Zooming-in and zooming-out appeared to recruit highly similar fronto-parietal networks, although the extent of activation was larger for zooming-in than zooming-out (Table 1b). Not only the extent of the activation clusters, but also the strength of activation was increased for zooming-in compared to zooming-out. This resulted from deconvolution analyses performed in each activation cluster of the rescaling network (Table 1a). Estimated hemodynamic response functions for the different conditions of a representative area (left aIPS) are shown in Figure 5a. As illustrated, this VOI showed a higher peak response when attention had to be zoomed-in compared to when attention had to be zoomed-out, a pattern which was also observed in the majority of the other VOIs (Figure 5b).

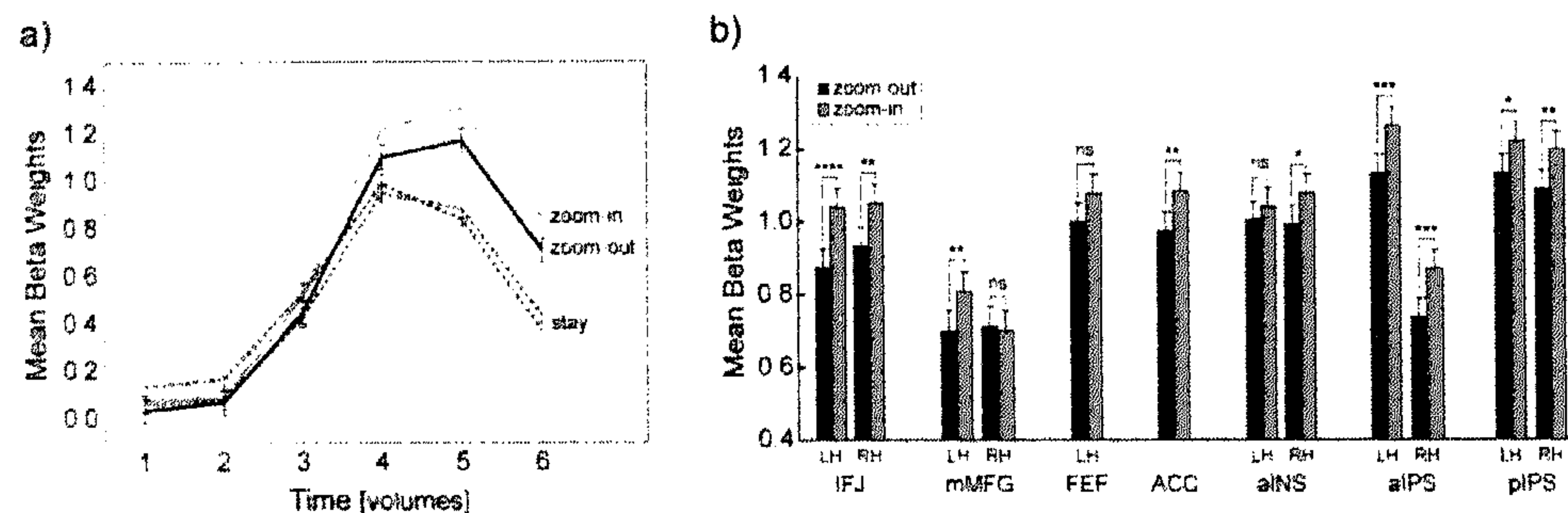


Figure 5. a). The estimated hemodynamic response function in the left lateral intraparietal sulcus is higher for the zoom-in compared to the zoom-out condition. Similar response patterns were obtained in the other VOIs. Each data point (i.e., each volume acquisition = 1.5 s) is independently estimated in the deconvolution analysis. The onset of the compound stimulus presentation was at data point 1. Error bars indicate SEM. b). Mean beta weights corresponding to the mean peak amplitude (i.e., the average of datapoints 4 and 5) for the zoom-in (light grey) and zoom-out (dark grey) condition obtained in the different activation clusters of the rescaling network (see Figure 4 and Table 1a for details). See text for abbreviations. Error bars indicate SEM. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.001$; **** $p < 0.0005$.

None of the activation clusters of the rescaling network showed a difference between stay-large and stay-small trials ($p > 0.1$) suggesting that the stronger responses in the zoom-in compared to the zoom-out condition were due to activation difference during the rescaling of the zoom-lens per se. In other words, differential responses appeared not to be confounded by initial differences in the size of the zoom-lens that were already present at the onset of the compound stimulus presentation.

Reflections of response incongruency costs in the dorsal anterior cingulate cortex

Deconvolution analyses were performed to estimate the differences between hemodynamic responses when the zoom-lens had to be rescaled to a target level that was response incongruent compared to congruent with the middle level. These analyses were confined to subject-specific VOIs (Table 2 & Figure 6a) in the

dorsal ACC, a brain region which is known to be involved in the monitoring for and detection of response conflict.

subject	left hemisphere				right hemisphere				t-value
	x	y	z	size	x	y	z	size	
1	-2	15	35	1190	8	15	32	1561	6.8
2	-	-	-	-	7	21	44	549	6.8
3	-7	29	32	383	7	24	41	563	4.0
4	-	-	-	-	6	14	39	364	4.0
5	-2	26	33	1619	9	29	28	873	6.8
6	-3	23	29	60	-	-	-	-	4.0
7	-4	5	47	750	4	12	52	601	8.0
8	-	-	-	-	4	18	43	959	6.8
mean	-4	20	35	800	6	19	40	781	5.9
sem	0.9	4.3	3.1	278	0.7	2.3	3.0	151	0.5

Table 2. Talairach coordinates (in mm) of the centre of gravity, the size (in mm³), and the t-values of the VOIs in the dorsal ACC identified for each participant. Mean and standard error of the mean (sem) are reported for identified VOIs in each hemisphere. Two participants were excluded from VOI analyses (one participant did not show dorsal ACC activation and the other had considerably slower and less accurate responses compared to other participants).

The average hemodynamic response in the dorsal ACC estimated in subject-specific VOIs (Figure 6a), peaked between 4.5 and 6 s after stimulus onset (i.e., data point 4-5 in Figure 6b), which is consistent with peak latencies observed in other event-related studies.

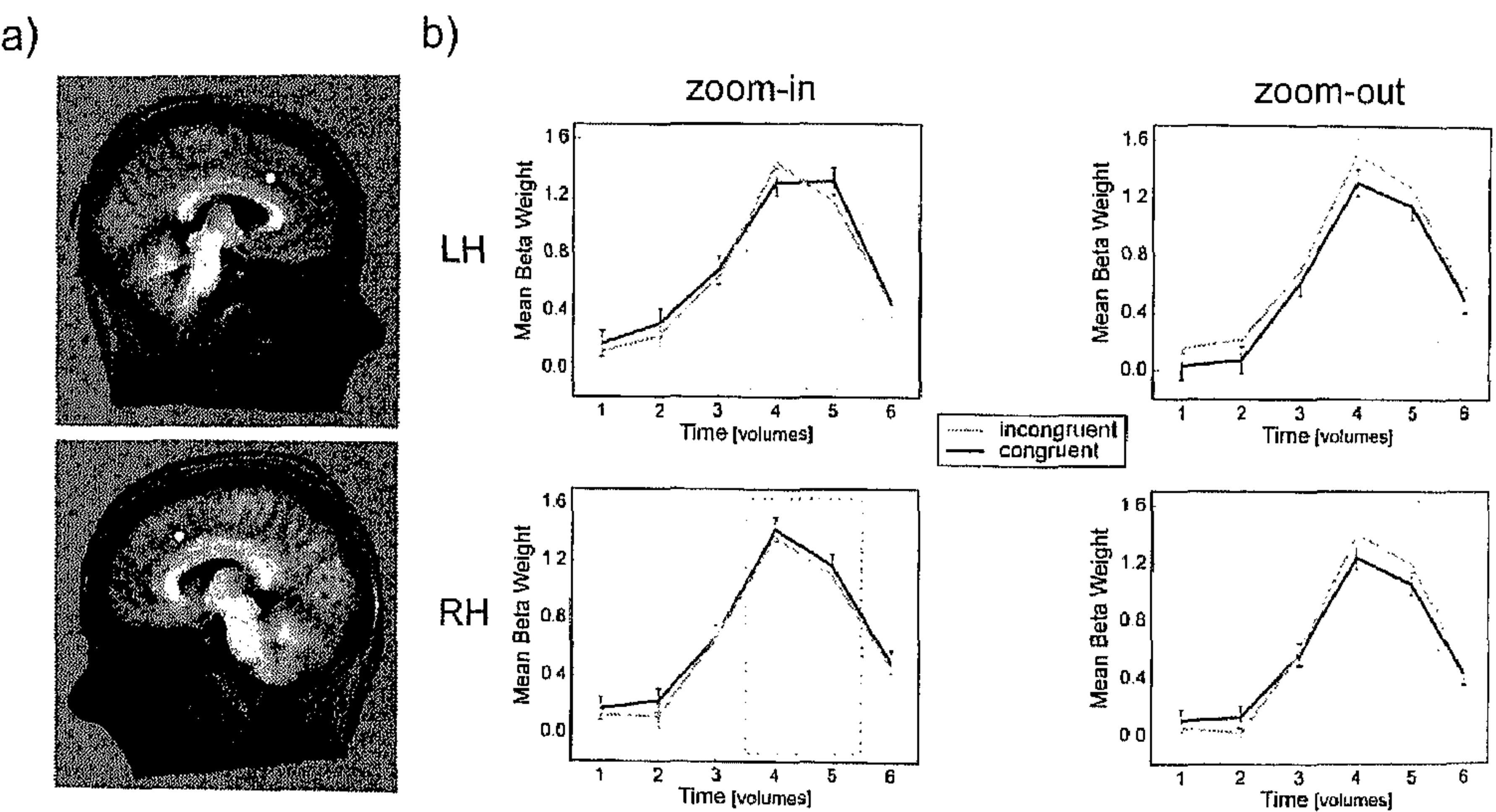


Figure 6. a) Location of the mean center of gravity of the subject-specific VOIs in the left (upper figure) and right (lower figure) dorsal ACC projected on the group average of the Talairach normalized high-resolution scans. b) Average hemodynamic response functions in the rostral cingulate zone estimated in subject-specific VOIs, shown in a), in the left (upper row) and in the right (lower row) hemisphere. Responses on zoom-out (right column) and zoom-in (left column) trials are separately shown for the condition in which the middle and target level were response congruent (dark colored lines) or response incongruent (bright colored lines). The onset of the compound stimulus presentation was at data point 1. The dashed boxes indicate the time window over which the contrast analysis was performed (i.e., the average of datapoint 4 and 5). Error bars indicate SEM. See text for further details.

In line with the behavioral response congruency costs, Figure 6b suggests that the amplitude of this peak is influenced by response congruency of the middle and target level in zoom-out trials, but not in zoom-in trials. These observations were confirmed by contrasting the beta weights corresponding to the peak latency (i.e., the average of the beta weights of data point 4 and 5) of the estimated response for the response congruent and incongruent condition for the zoom-in and zoom-out condition separately. The peak amplitude was higher in the incongruent compared to congruent condition when attention was zoomed-out (left: $t(4) = 1.8$; $p = .036$; right hemisphere: $t(6) = 2.0$; $p = .026$), but not when attention was zoomed-in (left: $t(4) = .07$; $p = .473$; right hemisphere: $t(6) = .07$; $p = .234$; all t-tests one-sided).

Discussion

The present study investigated the neural mechanisms of rescaling attention in a newly developed three-level compound stimulus. By comparing conditions in which attention could be held at the cued level (stay-trials) to conditions in which attention had to be rescaled from the cued level to the opposite level (zoom-trials), we obtained a fronto-parietal network that was specifically recruited when attention had to be rescaled. Moreover, the amount of attention that was deployed to the middle level during rescaling was studied by comparing the influence of response congruency between the middle and target level on activity in brain areas sensitive to response conflict.

Fronto-parietal network involved in rescaling the zoom-lens

Rescaling the zoom-lens from the cued to the opposite level of the compound stimulus, as opposed to holding attention on the cued level, led to enhanced activity in a fronto-parietal network. Frontal activations included bilateral areas in the middle part of the MFG, the IFJ, the anterior insula, the dorsal ACC, and the left FEF. In addition, large patches of bilateral activation were found in the IPS that extended into the SPL, PrCu and TOS. This network involved in rescaling attention overlapped with the fronto-parietal network recruited when attention has to be covertly shifted from one location to another (e.g., Beauchamp et al., 2001; Corbetta, et al., 2000; Gitelman et al., 1999; Hopfinger et al., 2000; Yantis et al., 2002). However, the studies that revealed this network involved in covert attention shifting, all used a paradigm in which the cue solely indicated the direction in which attention had to be shifted to detect the target that always had the same size. To our knowledge, the size of the cued area was varied over trials in only one fMRI study (Müller et al., 2003b). Therefore, in that study the attentional window did not only have to be shifted to the target location, but also had to be enlarged to encompass the entire region in which the target could appear. Shifting and scaling of the attentional window recruited a fronto-parietal network, which showed strong resemblance with the network obtained in the present study. Similar

activations were, among others, obtained in the aIPS, pIPS, the FEF, the IFJ⁶, and the ACC. Although these areas are also recruited when attention has to be shifted through space (e.g., Gitelman et al., 1999), the results of Müller et al. (2003b) revealed that activity in the right IFJ and left pIPS increased with the size of the cued area, suggesting that these areas were involved in expanding the attentional zoom-lens. The present study investigated (in isolation of attentional shifting processes) the influence of the direction rather than the extent of attentional scaling on neural activity. Highly similar fronto-parietal networks were involved in zooming-in and zooming-out of attention, although the extent of activation was larger for zooming-in. Moreover, zooming-in compared to zooming-out increased activity in most of the areas in the obtained rescaling network, including bilateral IFJ, dorsal ACC, IPS, and left DLPFC. Activation of these areas has been observed in tasks that require a high degree of cognitive or attentional control. For example, the dorsal ACC monitors for the presence of response conflict and signals the need for increased control to resolve detected response conflicts (Botvinick, Cohen, & Carter, 2004; Kerns, Cohen, MacDonald, Cho, Stenger, & Carter, 2004; MacDonald, Cohen, Stenger, & Carter, 2000). In turn, the DLPFC is recruited when the presence of conflicting distractors necessitate a strong attentional focus on task-relevant stimuli (Banich et al., 2000; MacDonald et al., 2000), which amplifies responses to target features in higher visual areas (Egner & Hirsch, 2005) possibly by influencing attentional processes in the parietal regions (Durstun et al., 2003). In addition, the IFJ appears to maintain task-relevant information that biases stimulus-response associations (Derrfuss, Brass, & von Cramon, 2004) represented in the posterior parietal cortex (Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002). Note that, since none of the rescaling network areas showed a difference between stay-large and stay-small trials, the initial size of the zoom-lens did not contribute to activation differences between expanding and contracting the zoom-lens.

In sum, these results showed that scaling of the attentional focus is mediated by a fronto-parietal network that closely resemble the network involved in shifting the attentional focus. Furthermore, zooming-in appeared to require a higher degree of attentional control than zooming-out, which could have been due to differences in the context of the small and large letter. That is, due to an increasing number of overlapping objects towards the middle of the stimulus, it might have been more difficult to segment the small letter from its background when zooming-in than segmenting the large letter from its background when zooming-out. Therefore, the influence of attention on figure-ground organization processes (Driver, Davis, Russell, Turratto, & Freeman, 2001; Vecera, Flevoris, & Filapek, 2004; Zenger, Braun, & Koch, 2000) might needed to be stronger in zooming-in compared to zooming-out, to accurately identify the presented letters.

⁶ Although the IFJ activation was labeled 'ventral precentral sulcus', its coordinates closely correspond to the activation in the present study.

Response congruency cost of the middle level reflected in the dorsal anterior cingulate cortex

In addition to studying the neural mechanisms involved in rescaling the zoom-lens, the present study also focused on neural indications supporting either a continuous or discrete rescaling of the zoom-lens. To investigate whether the zoom-lens temporarily encompassed the intermediate level during rescaling, response congruency between the different levels of the compound stimulus was varied. Behavioral performance was deteriorated when attention had to be rescaled from the small to the large level in a stimulus with an incongruent (compared to congruent) middle and target level. This behavioral response congruency cost for zooming-out in an incongruent stimulus was reflected at a neural level by increased dorsal ACC activity signaling the detection of response conflict (e.g., Botvinick et al., 2004). The presence of response conflict in incongruent stimuli suggested that the middle level apparently cannot be “skipped” when attention is rescaled from the small to the large level. Based on the training session, in which behavioral response congruency costs were not only obtained for zooming-out but also for zooming-in, we expected that the middle level would also interfere when attention was zoomed-in from the large to the small level. However, neither the behavioral nor the fMRI data suggested an interference of the middle level when attention was zoomed-in. Although interpretations concerning null-effects should be regarded with caution as they might be caused by insufficient statistical power, this lack of interference suggested that attention could be discretely rescaled while zooming-in (or at least in a more discrete manner than during zooming-out). The indications for discrete zooming-in and continuous zooming-out hinted at a direction-dependent fashion of rescaling. As suggested by the diverging patterns in rescaling performance in the different psychophysical studies discussed in Chapter 4, the exact manner of rescaling appears to be dynamically influenced by top-down factors, such as strategy and practice, as well as bottom-up factors, like visual context. This is in line with studies showing that attentional control can be adjusted by factors like strategy, even on a trial-by-trial basis (e.g., Egner & Hirsch, 2005). Moreover, recent fMRI studies suggested that response conflict between a target and a distractor increases during an attentional lapse (Weissman, Roberts, Visscher, & Woldorff, 2006) and can be reduced by increased attentional control (Egner & Hirsch, 2005). Speculatively, the increased attentional control when attention is zoomed-in compared to zoomed-out (see previous section) might have made an optimal zooming-in from the large to the small level possible (but see Chapter 4). In addition, the increased experience might have contributed to a more efficient zooming-in in the fMRI session compared to the training session. However, further studies are needed to uncover which factors contributed to the differential findings for zooming-in and zooming-out in the present study. Finally, one potential influence deserves mention. Similar to many other studies (e.g., Eriksen & Eriksen, 1974; LaBerge, Brown, Carter, Bash, & Hartley, 1991; McMains & Somers, 2004, 2005; Müller & Hübner, 2002), a letter identification

task was used to investigate the behavior of spatial attention. Although the results are discussed in the context of scaling spatial attention, object-based attention (e.g., Duncan, 1984; Vecera & Farah, 1994; for a review see Scholl, 2001) probably guided the rescaling of the zoom-lens, resulting in an attentional distribution that was most likely an interaction (e.g., Abrams & Law, 2000; Müller & Kleinschmidt, 2003) between spatial attention and object-based attention. However, since control of object-based and space-based attention shifting are mediated by highly similar networks (cf. Serences, Schwarzbach, Golay, Courtney, & Yantis, 2004; Yantis et al., 2002), we would expect no fundamental changes in the reported results if object-based attention could have been completely eliminated from the current design.

Conclusion

The present fMRI study investigated the neural mechanisms of rescaling the attentional zoom-lens using a three-level compound stimulus. Rescaling the zoom-lens recruited a fronto-parietal network, which largely overlapped with the network previously shown to be involved in shifting attention from one location to another. Zooming-in and zooming-out of attention recruited highly similar fronto-parietal networks, although the extent of activation was larger for zooming-in. In addition, areas within the rescaling network, thought to be recruited when an increased cognitive control is necessary, showed enhanced activity when attention had to be zoomed-in compared to zoomed-out. This finding could, however, have been influenced by differences in figure-ground segmentation difficulty between the large and small level. Furthermore, the fashion in which attention was rescaled also appeared to be influenced by the direction of rescaling. Activity in the dorsal ACC was enhanced, when attention was zoomed-out in a stimulus with an incongruent intermediate and large letter, indicating that the zoom-lens could not “jump” from the small to the large level. Conversely, no indications were found that the middle level interfered when attention was zoomed-in. In sum, the results suggested that although zooming-in and zooming-out recruited a highly similar fronto-parietal network, attention could be discretely rescaled from the large to the small level but not the other way around.

Acknowledgments

We acknowledge Dirk Vorberg for developing the stimuli and Joel Reithler and Bettina Sorger for comments on the manuscript. This research was supported by NWO Grant No. 402-01-632 to both authors.

References

- Abrams, R. A., & Law, M. B. (2000). Object-based visual attention with endogenous orienting. *Perception & Psychophysics*, 62(4), 818-833.
- Banich, M. T., Milham, M. P., Atchley, R. A., Cohen, N. J., Webb, A., Wszalek, T., Kramer, A. F., Liang, Z., Barad, V., Gullett, D., Shah, C., & Brown, C. (2000). Prefrontal regions play a predominant role in imposing an attentional 'set': evidence from fMRI. *Cognitive Brain Research*, 10(1-2), 1-9.
- Beauchamp, M. S., Petit, L., Ellmore, T. M., Ingeholm, J., & Haxby, J. V. (2001). A parametric fMRI study of overt and covert shifts of visuospatial attention. *Neuroimage*, 14(2), 310-321.
- Blanke, O., Spinelli, L., Thut, G., Michel, C. M., Perrig, S., Landis, T., & Seeck, M. (2000). Location of the human frontal eye field as defined by electrical cortical stimulation: anatomical, functional and electrophysiological characteristics. *Neuroreport*, 11(9), 1907-1913.
- Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in cognitive sciences*, 8(12), 539-546.
- Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, 402(6758), 179-181.
- Bunge, S. A., Hazeltine, E., Scanlon, M. D., Rosen, A. C., & Gabrieli, J. D. (2002). Dissociable contributions of prefrontal and parietal cortices to response selection. *Neuroimage*, 17(3), 1562-1571.
- Castiello, U., & Umiltà, C. (1990). Size of the attentional focus and efficiency components of processing. *Acta Psychologica*, 73, 195-209.
- Chastain, G. (1992a). Analog versus discrete shifts of attention across the visual field. *Psychological Research*, 54, 175-181.
- Chastain, G. (1992b). Time-course of sensitivity changes as attention shifts to an unpredictable location. *Journal of General Psychology*, 119, 105-111.
- Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, 3(3), 292-297.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, 3(3), 201-215.
- Derrfuss, J., Brass, M., & von Cramon, D. Y. (2004). Cognitive control in the posterior frontolateral cortex: evidence from common activations in task coordination, interference control, and working memory. *Neuroimage*, 23, 604-612.
- Driver, J., Davis, G., Russell, C., Turratto, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition*, 80, 61-95.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 501-517.
- Durston, S., Davidson, M. C., Thomas, K. M., Worden, M. S., Tottenham, N., Martinez, A., Watts, R., Ulug, A. M., & Casey, B. J. (2003). Parametric manipulation of conflict and response competition using rapid mixed-trial event-related fMRI. *Neuroimage*, 20(4), 2135-2141.
- Egner T., & Hirsch J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, 8(12), 1784-1790.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 143-149.
- Eriksen, C. W., & Hoffman, J. E. (1974). Selective attention: Noise suppression or signal enhancement? *Bulletin of the Psychonomic Society*, 4, 587-589.
- Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: a zoom lens model. *Perception & Psychophysics*, 40:225-240.

- Gitelman, D. R., Nobre, A. C., Parrish, T. B., LaBar, K. S., Kim, Y. H., Meyer, J. R., & Mesulam, M. (1999). A large-scale distributed network for covert spatial attention: further anatomical delineation based on stringent behavioural and cognitive controls. *Brain*, 122, 1093-1106.
- Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage*, 9(4), 416-429.
- Greenwood, P. M., & Parasuraman, R. (1999). Scale of attentional focus in visual search. *Perception & Psychophysics*, 61(5), 837-859.
- Hopf, J.-M., Luck, S. J., Boelmans, K., Schoenfeld, M. A., Boehler, C. N., Rieger, J., & Heinze, H.-J. (2006). The neural site of attention matches the spatial scale of perception. *The Journal of Neuroscience*, 26(13), 3532-3540.
- Hopfinger, J. B., Buonocore, M. H., & Magnun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience*, 3(3), 284-291.
- Kerns, J. G., Cohen, J. D., MacDonald III, A. W., Cho, R. Y., Stenger, V. A., & Carter, S. C. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, 303, 1023-1026.
- LaBerge, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 371-379.
- LaBerge, D., Brown, V., Carter, M., Bash, D., & Hartley, A. (1991). Reducing the effects of adjacent distractors by narrowing attention. *Journal of Experimental Psychology: Human Perception and Performance*, 17(1), 65-76.
- Luo, Y.-J., Greenwood, P. M., & Parasuraman, R. (2001). Dynamics of the spatial scale of visual attention revealed by brain event-related potentials. *Cognitive Brain Research*, 12, 371-381.
- MacDonald III, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288, 1835-1838.
- McMains, S. A., & Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron*, 42(4), 677-686.
- McMains, S. A., & Somers, D. C. (2005). Processing efficiency of divided spatial attention mechanisms in human visual cortex. *The Journal of Neuroscience*, 25(41), 9444-9448.
- Müller, M. M., & Hübner, R. (2002). Can the attentional spotlight be shaped like a doughnut? Evidence from steady state visual evoked potentials. *Psychological Science*, 13, 119-124.
- Müller, N. G., Bartelt, O. A., Donner, T. H., Villringer, A., & Brandt, S. A. (2003a). A physiological correlate of the "zoom lens" of visual attention. *The Journal of Neuroscience*, 23(9), 3561-3565.
- Müller, N. G., Donner, T. H., Bartelt, O. A., Brandt, S. A., Villringer, A., & Kleinschmidt, A. (2003b). The functional neuroanatomy of visual conjunction search: a parametric fMRI study. *Neuroimage*, 20(3), 1578-1590.
- Müller, N. G., & Kleinschmidt, A. (2003). Dynamic Interaction of Object- and Space-Based Attention in Retinotopic Visual Areas. *The Journal of Neuroscience*, 23(30), 9812-9816.
- Murphy, T. D., & Eriksen, C. W. (1987). Temporal changes in the distribution of attention in the visual field in response to precues. *Perception & Psychophysics*, 42(6), 576-586.
- Picard, N., & Strick, P. L. (2001). Imaging the premotor areas. *Current Opinion in Neurobiology*, 11, 663-672.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32A, 3-25.
- Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, 306, 443-447.
- Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition*, 80, 1-46.
- Serences, J. T. (2004). A comparison of methods for characterizing the event-related BOLD timeseries in rapid fMRI. *Neuroimage*, 21(4), 1690-1700.
- Serences, J. T., Schwarzbach, J., Golay, X., Courtney, S. M., & Yantis, S. (2004). Control of object-based attention in human cortex. *Cereb. Cortex*, 14, 1346-1357.

- Shulman, G. L., Remington, R. W., & MacLean, J. P. (1979). Moving attention through physical space. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 522-526.
- Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review*, 102(3), 503-532.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. Stuttgart: G. Thieme.
- Tsal, Y. (1983). Movements of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 523-530.
- Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, 123, 146-160.
- Vecera, S. P., Flevaris, A. V. & Filapek, J. C. (2004). Exogenous spatial attention influences figure-ground assignment. *Psychological Science*, 15(1), 20-26.
- Weissman, D. H., Roberts, K. C., Visscher, K. M., & Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nature Neuroscience*, 9(7), 971-978.
- Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., & Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience*, 5(10), 995-1002.
- Zenger, B., Braun, J., & Koch, C. (2000). Attentional effects on contrast detection in the presence of surround masks. *Vision Research*, 40, 3717-3724.

Chapter 6

Visual attention to multiple stimuli: To shift or to divide?

Based on: Jans, B., Peters, J. C., & De Weerd, P. Visual attention to multiple stimuli: To shift or to divide? (*submitted*)

Abstract

Despite a long tradition of attention research, the contribution of serial and parallel processes to visual attentional performance remains a matter of debate. This review offers an overview of studies that have used divided attention paradigms as a tool to investigate this issue. Studies of divided attention have focused on the question whether processing of several simultaneously presented targets can be truly parallel, without serial movements of focal attention. The combined evidence from psychophysical, electrophysiological, modeling and fMRI studies suggests that parallel models provide a plausible account of performance in divided attention tasks. Furthermore, the serial attention shifts that can be induced in specific paradigms take at least 100ms, making it biologically plausible to accommodate them in a flexible parallel model. We conclude that the distinction between serial and parallel models is useful at a descriptive level, but that an attention mechanism with a moderately flexible, parallel architecture is sufficient to accommodate both serial-looking and parallel-looking attentional behavior.

Introduction

Visual attention has limited capacity, and the question how a limited amount of attention is divided over multiple objects or locations in complex visual scenes has been at the core of attention research for almost two centuries. Attentional processing has mostly been studied in visual displays showing a single, relevant object (target), other objects being distracters. However, in today's society, the visual environment can be very complex, and optimal functioning may require the simultaneous processing of more than one source of information. For example, for pilots or other operators of complex equipment it can be beneficial to simultaneously process information from multiple measurement devices dispersed in the visual field. The question arises therefore how attention is divided over multiple sources of information.

This issue goes back to the old theories of early (Broadbent, 1954) and late (Deutsch & Deutsch, 1963) attentional selection. According to early selection theory, attentional enhancement of processing can occur only within a single window (the focus of attention), and selection takes place at an early stage of processing. In this view, the scanning of multiple stimuli is time-consuming because it requires moving the attentional focus. By contrast, according to late selection theory all objects in the visual field are processed in parallel at a speed that is determined by the division of attentional resources over all objects, and attentional selection of the target occurs at a late stage of processing. Physiological (Motter, 1994; Reynolds, Chelazzi & Desimone, 1999; Hopf, Luck, Girelli, Hagner, Mangun, Scheich & Heinze, 2000) and fMRI studies (Kastner, De Weerd, Ungerleider & Desimone, 1998; Ress & Heeger, 2003) indicate that attentional selection takes place at the earliest levels of visual sensory processing, but whether attention is allocated in parallel or serially has remained a matter of debate.

The coupling of attention with the foveating of objects almost naturally leads to the concept of a focus of attention, which 'illuminates' relevant objects in a scene (Eriksen & Eriksen, 1974). In their classical spatial cueing experiments, Posner and colleagues (Posner et al. 1980; Posner & Cohen, 1984), found shorter reaction times to a single stimulus when the position of that stimulus was cued beforehand. These spatial cueing data were interpreted as evidence for a serially moving focus of attention.

Search paradigms may be better suited than Posner's paradigm to evaluate the contribution of serial and parallel processes to attentional selection, because of the necessity to allocate attention to multiple items. In search tasks, target search times independent of the number of distracters have been interpreted as evidence for parallel processing, while search times that increase with the number of distracters have been interpreted as evidence for a serial process. Typical serial search tasks show a linear relationship between reaction time and number of distracters in target-present displays (20-100ms added per distracter), and show a 2:1 ratio between the slopes of search curves in target-present and target-absent

displays (e.g., Estes & Wessel, 1966; Eriksen & Spencer, 1969; Treisman & Gelade, 1980). Because the above data suggested a dichotomy between serial and parallel processes, a number of early studies have focused on determining the stimulus properties that would trigger serial or parallel search (e.g., Treisman & Gelade, 1980; Bergen & Julesz, 1983). Later studies have produced support for mixed models of attention (Shiffrin & Schneider, 1977; Duncan, 1984; Treisman & Gormican, 1988; Krose & Julesz, 1989; Wolfe, 1994). In these models, parallel processing (sometimes referred to as pre-attentive) of items across the visual field is followed by a serial process that leads to target selection. A recent neurophysiological study has provided direct support for a mixed model. In this elegant study (Bichot, Rossi & Desimone, 2005), monkeys performed a conjunction search task while freely viewing the search display. Bichot and colleagues demonstrated that during the scanning of the display with eye movements, there was a parallel enhancement of activity in neurons with receptive fields (RFs) away from the fovea that contained an item sharing a feature with the search target. The strongest enhancement was observed when the item in the RF was a target and when it was selected for an eye movement. In this study, parallel processes appeared to isolate a number of candidate targets, which were inspected serially by eye movements until the target was found. Note, however, that the 'serial-looking' focusing of attention on candidate targets during target selection may reflect a flexible parallel distribution of attention (Bundesen, 1990), rather than serial scanning by an inflexible focus of attention.

Divided attention paradigms in which two (or maximally a few) objects must be detected or discriminated simultaneously have been another approach to test serial and parallel models of attention. The *serial models* that have been used to explain performance in divided attention tasks are of two types. According to a *serial shifting theory of attention* (Eriksen & Eriksen, 1974; Posner, 1980), (quasi-) simultaneous processing of two objects is related to rapid shifting of a single, indivisible focus of attention that is fixed in size. According to a *zoom-lens theory of attention* (Eriksen & James, 1986), the attentional focus can be 'zoomed' to match the spatial scale of the attended object. Zooming the focus of attention implies parallel processing within the zoomed focus of attention, and could be used to explain simultaneous processing of two separate target stimuli falling within the zoomed focus of attention. Thus, serial models could accommodate data suggestive of parallel processing in two ways. They might either assume *rapid shifting* of the attentional focus between separate stimulus locations (*scanning*), or *zooming* of the focus of attention so it can encompass multiple information sources in the visual field.

To demonstrate that rapid serial scanning is not necessary for accurate performance in divided attention tasks, many studies have used short stimulus presentations. Estimates of how short stimulus presentations have to be to prevent attention shifts depend somewhat on the specific paradigm used. Several exogenous cueing studies have shown that after directing attention to a primary

location with a physical cue, identification of a target appearing in a non-cued location delays target identification by 110-160ms compared to identification in the cued location (Eriksen & Yeh, 1985; Krose & Julesz, 1989), or more if stimuli are masked (Krose & Julesz, 1989). In difficult serial search tasks that require individual scrutiny of items in the search array, search times increase by roughly 100ms per item (Bricolo, Giansini, Fanini, Bundesen & Chelazzi, 2002). Other studies using the 'attentional dwell paradigm' (in which a stimulus is presented and masked in one location, and followed by a stimulus and mask in a second location) have shown that the inter-stimulus interval required to identify both stimuli exceeds 200ms (Krose & Julesz, 1989; Duncan, Ward, & Shapiro, 1994; Moore, Egeth, Berglan, & Luck, 1996; Ward, Duncan, & Shapiro, 1996). The behavioral data taken together suggest that shifts in focused attention between two items/locations cannot be achieved in less than 100ms. Electrophysiological studies in monkeys (Motter, 1994; Khayat, Spekreijse & Roelfseman, 2006) and human subjects (Mangun, 1995; Müller, Teder-Salejarvi & Hillyard, 1998; Hillyard & Anllo-Venlo, 1998; Woodman & Luck, 1999, 2003) indicate that attention switches require 150-200ms. Thus, if two target stimuli in a display are presented simultaneously, stimulus durations below 200ms strongly reduce the probability that a serial mechanism can account for their successful identification, on the condition that the identification of the target stimuli is attention-demanding.

Both the rapid shifting hypothesis and the zooming hypothesis predict attentional enhancement in-between stimulus locations, and this enhancement can be revealed with probe trials. Under the zooming hypothesis, the attentional enhancement would apply to the entire focus of attention and include, therefore, also regions in-between relevant locations. Under the rapid switching hypothesis, attentional enhancement in-between relevant locations is expected because the focus of attention is assumed to remain 'on' as it moves between relevant stimulus locations (for review Eriksen & Murphy, 1987). Thus, both variants of the single-focus, serial model of attention can be falsified by showing the absence of attentional modulation in-between relevant locations.

Truly *parallel processing* of information in different, non-adjacent regions in the visual field requires that one leaves behind the idea of a unitary focus of attention. Instead, multiple regions of attentional enhancement would exist, in the absence of enhanced attentional processing in-between these regions (*multiple foci theory of attention, or splitting model*).

Could a serial shifting hypothesis also be an alternative for a divided attention hypothesis in the absence of attentional modulation in-between stimulus locations? This merely requires the focus of attention to be turned 'off' in-between during its movement. In this way, for the most 'parallel' of findings, there would be a 'serial' explanation. The evidence for turning off the focus of attention during attention shifts is mixed, however, (Shulman, Remington & McLean, 1979; Tsal, 1983; Briand & Klein, 1987; Chastain, 1992; for review Eriksen & Murphy, 1987), and in the absence of compelling evidence it is more parsimonious to

assume that the attention focus remains ‘on’ while it moves (see also Discussion). The strongest evidence against serial models, however, would come from studies that minimize the possibility for serial shifts between relevant locations/stimuli by limiting exposure duration, and demonstrate the absence of attentional modulation in-between attended locations.

The specific goal of this review is to evaluate studies claiming to have created conditions under which the processing of multiple target stimuli (usually two) was either strictly parallel, or strictly serial. This review deals primarily with the division of attention between targets presented in different locations, and only briefly touches upon issues related to the division of attention among features within a single object (e.g., see Duncan, 1984; Driver & Baylis, 1989). Psychophysical studies on divided attention (e.g., Shaw & Shaw, 1977; Shaw, 1978; Egly & Homa, 1984; Müller & Findlay, 1987) often provided the inspiration for the paradigms and stimuli used in neuro-cognitive studies. Psychophysical studies therefore will be discussed first, followed by electrophysiological, imaging and computational studies.

Psychophysical studies

In psychophysical studies, the use of a serial or parallel processing strategy is derived indirectly from reaction time and accuracy data. An assumption in studies of parallel processing of two (or more) information sources is that, in the absence of biases or differences in processing requirements, a fixed amount of attentional resources is equally divided between attended locations, irrespective of their size. This implies that per unit of space, more resources will be available in a smaller attended location than in a larger one. Sensory processing of stimuli placed in the smaller region, therefore, can be enhanced more by attention than stimuli placed in a larger region (Eriksen & James, 1986; Castiello & Umiltà, 1992). Thus, according to a parallel model, stimuli placed in the smaller of two attended regions will be processed fastest. According to a serial model, however, the faster processing in a smaller attended location might reflect a faster completion of a serial scanning process (e.g. if the scanning speed and the size of the focus of attention were fixed).

Cueing of two locations by a single bar

Eriksen and Yeh (1985) obtained reaction time data that at first sight suggested a strong bias towards serial processing when subjects had to monitor information in several relevant locations. In their main experiment (Figure 1), 8 small stimuli were displayed together at a given eccentricity (2.5° in their ‘large’ displays) around a fixation spot, and the task of the subjects consisted of recognizing a target letter (S or Y). The target could appear in four possible locations: at 3, 6, 9 or 12 o’clock; stimuli in the other locations were distracters (Figure 1).

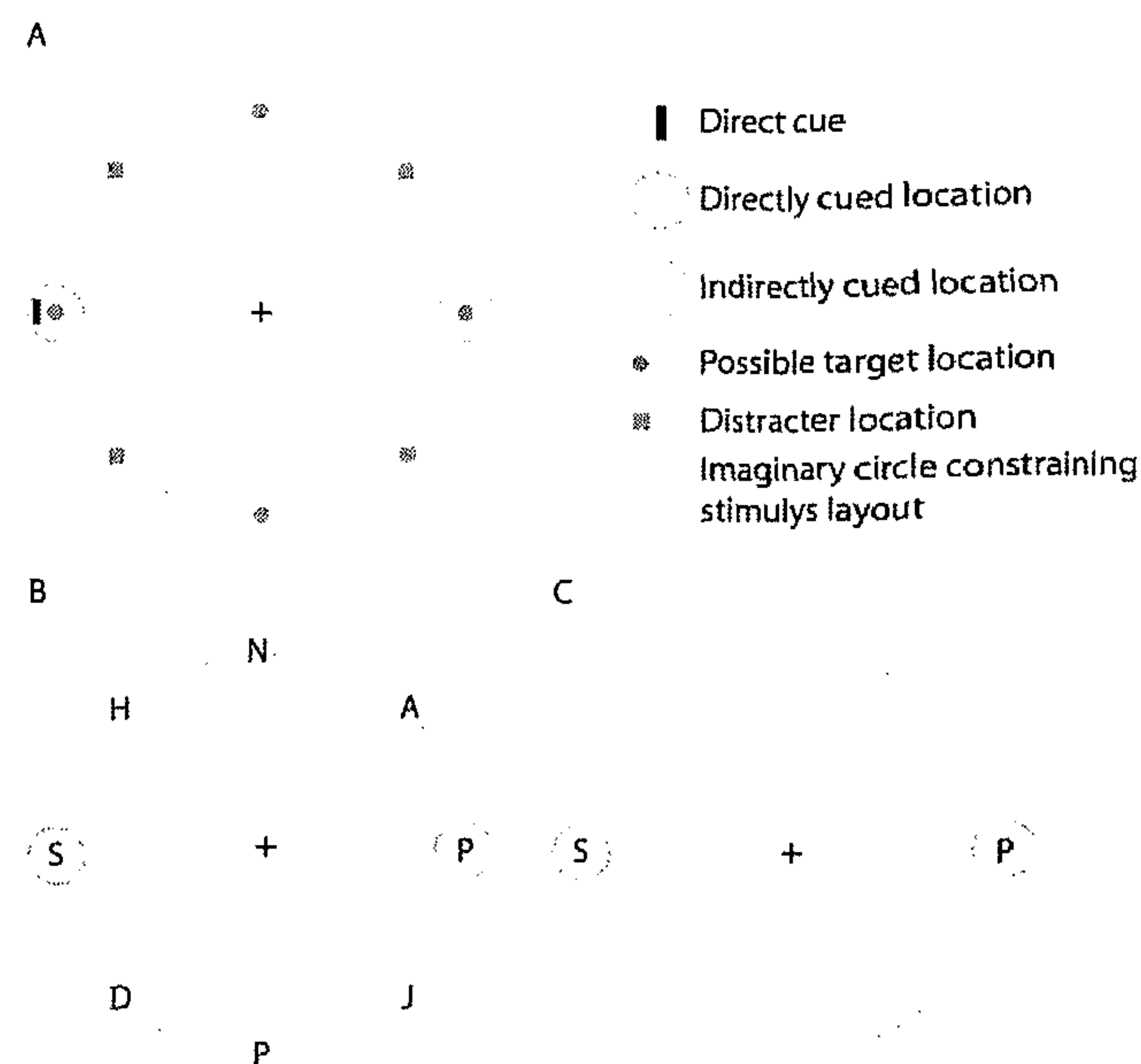


Figure 1. Overview of stimulus displays (not to scale) in the experiment of Eriksen and Yeh (1985). Light grey circles, dots and squares are for illustrative purposes, and were not present in the stimulus displays. In 'large displays', eccentricity of the stimuli was 2.5° ; in 'small displays' eccentricity was 1° . Results from both display sizes were indistinguishable and pooled by the authors. The width and height of the characters was 0.16° and 0.20° , respectively, and a similarly sized bar (0.20° by 0.08°) was used as a cue. A) A maximum of 8 stimuli as indicated by light grey dots and squares was presented in a given display, equi-eccentrically around fixation (dashed circle). There were 4 pre-defined potential target locations (illustrated by small dots), and 4 distracter locations (small squares). In any display, there were only two locations that could contain a target (targets were S or Y symbols). One location was cued directly by a small bar stimulus (small solid circle outline), the other indirectly by virtue of being opposite to the primary cue (small dashed circle outline). B) Example of a display with 8 stimuli, with a target (S) in the directly cued location, and a distracter (P) in the indirectly cued location. C) Example of a similar display with stimuli only in the target locations. For further details, see text.

In the experimental condition, 150 milliseconds (ms) before the onset of the display, a single bar-shaped cueing stimulus was presented near one of the four possible target locations. This cue not only indicated that a target was likely to occur in that location (the primary location), but also in the diametrically opposite location (secondary location). The probability that a stimulus would occur in the primary, the secondary or the two non-cued target locations was varied in three different experimental conditions. In the first condition (40-40-10-10), the target appeared on 40% of the trials in the primary location, on 40% of the trials in the secondary location and on 10% of the trials in the two remaining (non-cued) target locations. In the second condition (70-10-10-10), the target appeared on 70% of the trials in the primary location and on 10% of the trials in each of the remaining three locations. In the third condition (100-0-0-0), the target always appeared in the primary location. Subjects were informed of these probabilities before the start of the experiment. Eye movements were not monitored. The authors hypothesized that subjects would adopt attentional strategies according to the probabilities of

target presentation in the target locations. They predicted that the first condition (40-40-10-10) would induce parallel processing, resulting in equal reaction times for both cued target positions. A serial strategy was anticipated in the 70-10-10-10 condition, in which the subjects were strongly biased to expect the target in the primary target location. That strategy would result in shorter reaction times for the primary location than for the secondary and non-cued locations. In contrast to these predictions, the data were in agreement with the serial alternative in both experimental conditions. Reaction times to letters in the primary target location were always shorter than reaction times to letters in secondary or non-cued target locations, with the longest reaction times found in the non-cued locations. In the 40-40-10-10 condition, responses to secondary or non-cued locations were delayed by about 54 and 107ms, respectively, compared to the primary location. In the 70-10-10-10 condition, responses to secondary or non-cued locations were delayed by about 84 and 121ms, respectively, compared to the primary location.

Eriksen and Yeh (1985) suggested that subjects had a preset serial strategy to attend the primary location first, and to attend the secondary location or non-cued location later, even in the 40-40-10-10 condition where a more parallel strategy had seemed more likely a priori. While the interpretation of the global pattern of data in terms of a shift of attention is reasonable, the interpretation of the data in the 40-40-10-10 condition as evidence in favor of a *preset* serial strategy is not. The investigators cued the primary location directly with a small bar stimulus preceding the target stimulus, and cued the secondary location only indirectly. The time between the cue and the stimulus onset was 150 ms, which makes the cued location subject to exogenous facilitation, at the disadvantage of the secondary location. Posner and Cohen (1984), for example, demonstrated that visual processing of a stimulus is facilitated by attention if it is preceded by a cue in the same (or adjacent) location maximally 0-200 ms before stimulus onset. Both neurophysiological studies in monkeys (Gottlieb, Kusunoki, & Goldberg 1998; Kusunoki, Gottlieb, & Goldberg, 2000) and psychophysical studies in humans (Jonides & Yantis, 1988; Theeuwes, 1995; Enns, Austen, Di Lollo, Rauschenberger, & Yantis, 2001; Donk & Theeuwes, 2001; Franconeri & Simons, 2005) indicate that exogenous cues such as sudden onset can automatically capture attention, often in spite of competing endogenous cues. Thus, also in the 40-40-10-10 condition where target probabilities were equal in primary and secondary location, the primary location was cued more effectively than the secondary location. The results from this experiment, therefore, do not represent conclusive evidence against the possibility of simultaneous allocation of attention to two segregated locations. Likely, the reported results reflect a mixture of the strategies adopted by the subjects under the influence of exogenous cueing effects, and biases induced by the proportions of targets at different locations.

Some evidence from Eriksen and Yeh (1985) does indeed point to a contribution of parallel processes. They repeated their experiment in the absence of distracters (Figure 1C), and replicated the finding that reaction times to targets

in the primary location were shorter than to targets in other locations. In contrast to the previous experiment, reaction times to targets appearing in the secondary target location were the same as reaction times to targets appearing in non-cued locations, exceeding reaction times in the primary location by 20-30ms. This held true in both the 40-40-10-10 and the 70-10-10-10 conditions. Taken in isolation, this suggests that allocation of attention to the primary target location deprives all other target locations from attentional resources, in agreement with a serial model. Combined with the previous experiment, however, the data might rather be interpreted in favor of a parallel model, because the presence or absence of distracters in the 4 non-target locations influenced the attentional allocation to the 4 target locations. The results indicate that there is attentional allocation outside the primary focus of attention, with a stronger competition between target locations and a stronger effect of biasing instructions the more distracter elements there are in the display. The suppression of distracters thus is an active process that spends attentional resources at the detriment of target processing.

Cueing of two locations with differently-sized frames

Castiello and Umiltà (1992) performed a series of experiments aimed at reinvestigating the possibility of parallel processing of two locations (eye fixation was monitored). The subjects' task was to respond as fast as possible to the onset of a small dot. To cue the relevant locations within which the dot could appear, two empty frames (outlines of squares) were presented together to the left and right of fixation. The sizes of the two frames were never the same and varied on different trials (Figure 2A). The target dot (100ms) appeared at random in one of the two frames, 500ms after frame-onset. The frames stayed on until the end of the trial. Subjects were instructed to signal the appearance of the target dot, which was presented at random in one of the two frames. Because of the 500ms delay between frame-onset and target-onset, exogenous cueing effects that could have interfered with a division of attention according to the size of the frames were unlikely. The data showed that reaction times were smaller for dots presented in the small frame compared to dots presented in the larger frame. This is in agreement with a parallel model in which a limited amount of attentional resources (Eriksen & St. James, 1986) is split equally between the two locations, leading to more attention per unit of surface in the smaller frame. A serial account of the data seems unlikely given the relatively brief target presentation (100ms), but cannot be excluded because the stimuli were not masked. Such serial account would involve a small focus of attention, requiring more time to scan a large frame compared to a small frame.

The study of Castiello and Umiltà (1992) described above was replicated by McCormick, Klein and Johnston (1998), but they modified the design of the experiment to assess whether there was attentional modulation in-between relevant locations (eye fixation was monitored). They adopted the idea of cueing two locations enclosed within frames and had subjects report the presence of dot

stimuli in or near the cued regions (Figure 2B). The target dot was presented 515ms after onset of the frame(s), and both frame(s) and target dot remained on screen until the subject responded. There were three types of trials: with a single frame on the left of fixation, with a single frame on the right (single frame conditions), or with a frame on either side of fixation (double frame condition). In all three types of trials, a target dot could be presented in 4 possible locations (falling on a horizontal line through the fixation target). In the single-frame conditions, the dot could appear within the frame, in-between the frame and the fixation spot, and in the two symmetrical locations on the other side of the fixation spot. In the double-frame condition, 2 of the target locations fell within the frames, and the 2 others fell in-between the fixation spot and the frames. In any single-frame trial, there was a 0.62 probability that the target was presented inside the frame, with a much smaller probability for the remaining 3 locations outside the frame. In any double-frame trial, there was a 0.75 probability that a target appeared within one or the other frame (0.375 for each frame), with a much smaller probability for the remaining two locations outside the frames. Subjects were aware of these contingencies. The different cueing conditions were designed to test whether increasing the number of primarily attended locations (cued by frames) from one to two would lead to an increase in size of the attention focus to encompass both (zooming model), or whether this would lead to two separate attention foci (splitting model). Both models predict that in the single-frame condition, reaction times to targets within the frame would be shorter than to targets outside the frame. However, in the double-frame condition, the two models make different predictions. According to the zooming model, reaction times to targets shown in-between the two frames should be similar to reaction times to targets within a frame (because in-between targets fall within the enlarged attention focus). According to the splitting model, reaction times to targets shown in-between the two frames should be larger than reaction times to targets within a frame (because in-between targets fall in-between the separate foci of attention). The results supported the zooming model. Targets shown within frames always had a higher eccentricity than targets in-between a frame location and the fixation spot, but control experiments indicated that this did not influence the results.

Despite the enhancement in-between cued locations, the data could be explained by a moving focus of attention, because the target and frame(s) remained on the screen until the subject responded. With reaction times in the order of 300-400ms, a focus of attention has ample time to move between the two frame locations. Furthermore, a fundamental question is whether a probability of 0.375 to find a target in one of the two relevant locations in the divided attention condition is sufficient to induce a split in the focus of attention.

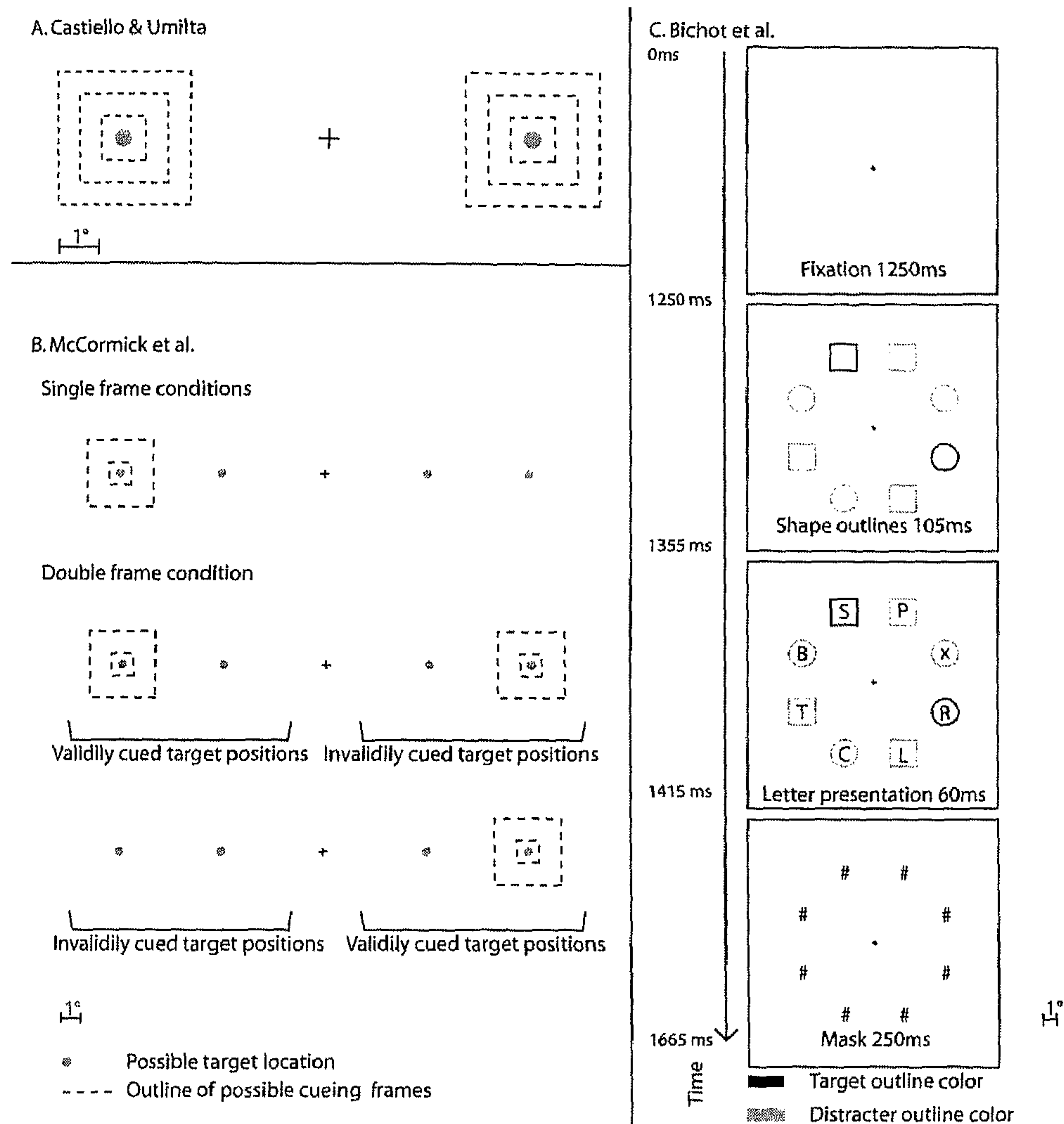


Figure 2. Overview of conditions in three experiments using frames as attentional cues (drawings to scale). A) Design and stimuli used by Castiello and Umiltà (1992). A single target could appear in one of two locations 5° to the left or right of fixation (light gray dots), both of which were cued simultaneously by frames (dashed lines) sized 1.1°, 2.2° or 3.3°. All sizes are illustrated, but actual trials only presented a single frame per position. B) Design and stimuli used by McCormick et al. (1998). A single target could appear in one of four locations 5° or 10° to the left or right of fixation (light gray dots). Two frame sizes were used (1.1° and 3.3°) in a double-frame condition, but also in a single-frame condition. In the latter condition, targets could be validly or invalidly cued by the preceding frames. C) Design and stimuli used by Bichot et al. (1999). Sequence of stimulus displays illustrated from top to bottom. After fixation (1250ms), 8 empty shape outlines (circles and squares) were presented (105ms) at an eccentricity of 6.1°. Subsequently, a letter was placed in each outline (60ms) and masked (250ms). Subjects compared the shapes of 2 target outlines that differed in color from the others (targets illustrated in black and distracters in grey), and also reported as many letters as possible.

Bichot, Cave and Pashler (1999) performed a location cueing experiment to further investigate the possibility of splitting the focus of attention. In each trial, 8

shape outlines were presented at 6.1° of eccentricity around a fixation spot, two of which were distinct in color and served as target regions (Figure 2C). The outlines were presented for 105ms, after which a letter was placed inside each outline for 60ms, upon which the display was masked. The subject had to report whether the two target-outlines matched in shape, and also had to report as many letters as possible. Eye fixation was not monitored. On the shape matching task, subjects reached a performance of 85% correct for adjacent shapes, with a small but significant drop for nonadjacent ones (80%). Within correctly discriminated outlines, subjects correctly identified up to 40% of the letters. Virtually no correct identifications took place, however, within distracter outlines, and this held true even when a single distracter outline was positioned in-between two target outlines. Bichot et al.'s (1999) data thus support a split focus model of attention.

In Bichot et al.'s (1999) cueing experiments, the absence of correct letter identifications in distracter regions between target locations is evidence against a zooming lens theory, and against serial movements of a focus of attention. The latter is also supported by the fact that letters were presented for only 60ms and then masked, and several studies have shown that in tasks that encourage shifts of focal attention, this cannot be done in less than about 110ms (see Introduction). The data thus support a split focus model of attention.

Despite the strong evidence of separate attentional foci, processing in the two foci may not have been fully independent. The probability of correctly identifying a target letter within a target region was about 0.4 (Figure 13 in Bichot et al. 1999). If this were the output of a truly parallel mechanism, working independently at the two target locations, then the probability of reporting two target letters correctly (both in a target region) on a single trial should be maximally $0.4 \times 0.4 = 0.16$. However, the authors reported a conditional probability in a 0.3-0.4 range that a target was identified correctly in one target region given the other target was identified correctly in the other target region (Figure 14 in Bichot et al., 1999). Although not explicitly tested by the authors this, this is probably significantly more than the 0.16 predicted by an independent, parallel mechanism. This stronger-than-expected correlation might reflect simultaneous (dependent) use of some resources at both locations. This could be expected if the source of the attentional enhancement within the separate foci is the same, which would lead to correlated variability in attentional processing in both locations.

Using masked cueing and prior entry to study division of attention between two locations

It has been shown that attention improves spatial resolution, thereby leading to improved accuracy of perception, increased discrimination performance, and shorter reaction times (Frey, 1990; Intriligator & Cavanagh, 2001; Shore, Spence, & Klein, 2001; Deco, Pollatos & Zihl, 2002; Schneider & Bavelier, 2003). In addition, attention has been demonstrated to speed up processing. Of two simultaneously presented stimuli, the onset of the attended stimulus will be

perceived as earlier than the non-attended one, a phenomenon termed ‘prior entry’ (Titchener, 1908). Prior entry can be induced by endogenous or exogenous cues. Exogenous cues can speed up processing if they are presented in sufficiently close spatial and temporal proximity of the target stimulus. When their spatial and temporal proximity is very close, the exogenous cue can become masked by the target (Breitmeyer & Kersey, 1981; Enns & DiLollo, 2000). Interestingly, in this case the cue does draw attention (McCormick, 1997; Jaskowski, van der Lubbe, Schlotterbeck & Verleger, 2002; Scharlau & Neumann, 2003a,b), although the conscious perception of the cue itself is strongly reduced or even absent due to masking (Klotz & Neumann, 1999; Scharlau & Ansorge, 2003). This ‘masked-cueing’ paradigm was used by Scharlau (2004) to investigate the division of attention between two non-contiguous locations in the visual field.

All three attention models proposed in the Introduction were contrasted with each other: (1) a *shifting model* in which a single focus of attention must be moved to scan multiple locations, (2) a *splitting model* in which attention is split into multiple foci in non-contiguous regions of the visual field, and (3) a *zooming model* in which a unitary focus can be expanded in size. A first experiment was mainly designed to falsify the shifting model, and a second experiment was designed to falsify the zooming model (Figure 3). In both experiments, subjects were asked to judge the order in which two differently colored disk outlines were presented (e.g., ‘red before green’ or ‘green before red’). These target stimuli were effectively presented with different onsets, the standard target (e.g., green) being presented at a fixed time in the course of the trial, and the variable target (e.g., red) being presented at a variable time interval before (negative SOA) or after the standard (positive SOA). The two target stimuli were presented together with a distracter stimulus (in a third color, e.g., yellow) in a triangular formation centered on the fixation spot (Figure 3). The distracter stimulus was always presented simultaneously with the variable target. Eccentricity of disk outlines (diameter 1.9°) was 5.6° , their position was randomized trial-by-trial, and different colors for targets and distracters were assigned to different subjects.

In the first experiment (Figure 3), four conditions were used, with the last one critical to falsify the unitary focus theory. First, in an *uncued condition*, the two targets were presented with different SOAs (-48 to $+48$ ms) without attentional manipulation. In this condition, the psychometric curve plotting the proportion responses ‘standard before variable’ against SOA would be unbiased, with 50% responses at a zero-SOA. Second, in a *single-cue condition* cueing the variable stimulus, prior entry would be induced for that stimulus, causing the psychometric curve to shift to the right (e.g., less than 50% ‘standard before variable’ response at a zero-SOA). This would be compatible with any of the three theories. Third, in a *target-target cueing* condition, the targets were both cued. All three theories would predict that the amount of prior entry induced in the standard and variable target stimuli would not differ, such that the proportion of responses ‘standard before variable’ would remain unaffected. According to the unitary focus model,

attention would be allocated at random to one or the other target, leading on average to the same amount of attention to both stimuli. Although the absolute amount of attention allocated to each target might be different in the zooming and split foci models, both would predict the same amount for each target. In a crucial, fourth condition (*target-distracter cueing*), two cues were used, one preceding the variable target, and another one preceding the distracter. According to the unitary focus theory, random assignment or shifting of the focus between the cued locations would lead to a reduction of the prior entry effect for the variable stimulus compared to the single cue condition. This would result in a smaller rightward shift of the psychometric distribution compared to the single cue condition. A similar result might be expected under a zooming hypothesis, assuming that attentional resources per unit of surface decline when the zoom lens expands. However, if two separate and equally potent foci of attention could be achieved, as predicted by the multiple foci theory, then the results in the condition where the variable target and the distracter were both cued *would not differ* from the results in the single-cue condition. Note that according to Scharlau (2004), the splitting of attention does not reduce the amount of attention available in individual foci, while Castiello and Umiltà (1992), and McCormick et al. (1998) assumed a division of attention between foci.

As expected, there was no difference between the no cue condition and the target-target cueing condition. More interestingly, the data showed that psychometric curves for the single cueing condition (cueing the variable target) and the variable-target/distracter cueing condition were significantly and equally shifted to the right, indicating an identical prior entry effect for the variable target in both conditions. The absolute prior entry effect was in the range of 30-40ms in both conditions. This is compatible with the splitting model, but not with the shifting model. In principle, the result could also be accommodated by a zooming theory assuming that the increase of the size of the focus would *not* go at the cost of decreased power.

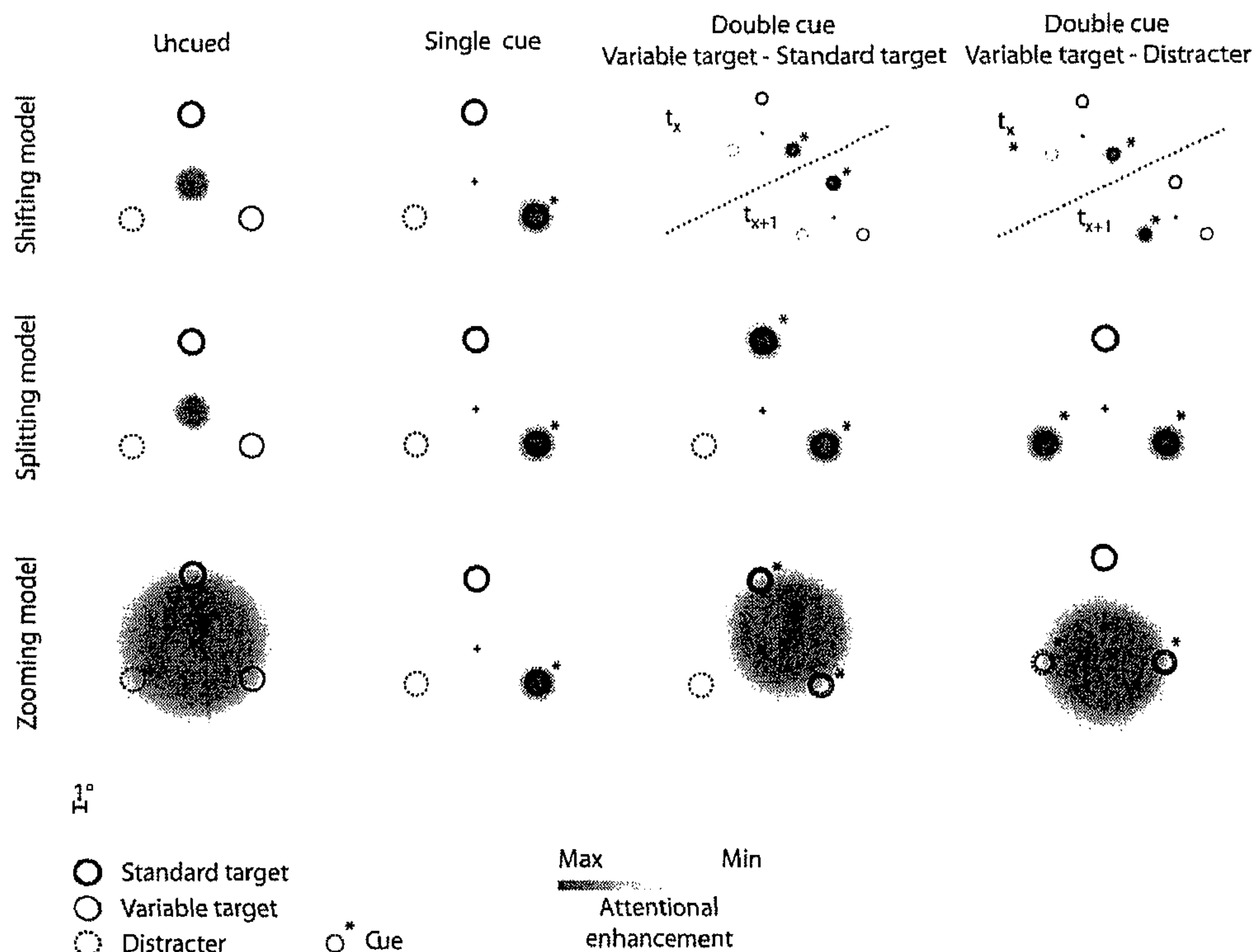


Figure 3. Experimental conditions in the prior-entry experiment of Scharlau (2004). Hypothesized shapes of attention distributions (radial grey scale gradients) are shown in four cueing conditions (4 columns) according to three attention theories (3 rows). Cue and stimulus diameters were 1.4° and 1.9° , respectively (drawing approximately to scale). For clarity, cue stimuli (small disk outlines) are indicated by asterisks. The movement of attention according to the shifting model is illustrated by showing the location of focused attention in the display at two time points, denoted t_x and t_{x+1} . For illustration, a fixed arrangement of standard, variable target and distracter was chosen, but in the experiment the position of the stimuli was randomized for the shown triangular configuration, and also a top-down triangular configuration was used. For further details, see text.

In the case of two separate, relevant stimuli, the zoom-lens model predicts that a unitary focus of attention is expanded to include both stimuli, leading to attentional facilitation at the locations in-between the target locations. To the contrary, the splitting model predicts an absence of attentional facilitation for the in-between locations. To falsify the zooming theory, a second experiment was carried out to test whether there was attentional enhancement of processing speed in-between the two targets (Figure 4). As in the previous experiment, there was an uncued condition, and a single target cueing condition in which the variable stimulus was cued. In the two other conditions the variable target stimulus and distracter were both cued (variable-target/distracter cueing). In one of these two conditions, the standard was placed away from the cued stimuli (*standard away*). In the other, crucial condition, the standard was placed in-between the cued stimuli (*standard in-between*). As discussed above, the zooming theory is compatible with a rightward shift of the psychometric curve in the single cue condition, and in the

standard away condition (relative to the psychometric curve in the no cueing condition). In the standard in-between condition, however, this theory predicts less of a rightward shift, because the standard stimulus would be cued together with the variable stimulus as they would fall together within the zoomed focus of attention. If one assumes that the distribution of attention within a focus is Gaussian, then one might even expect a stronger prior entry effect for the standard than for the variable target, which would result in a leftward shift of the psychometric distribution.

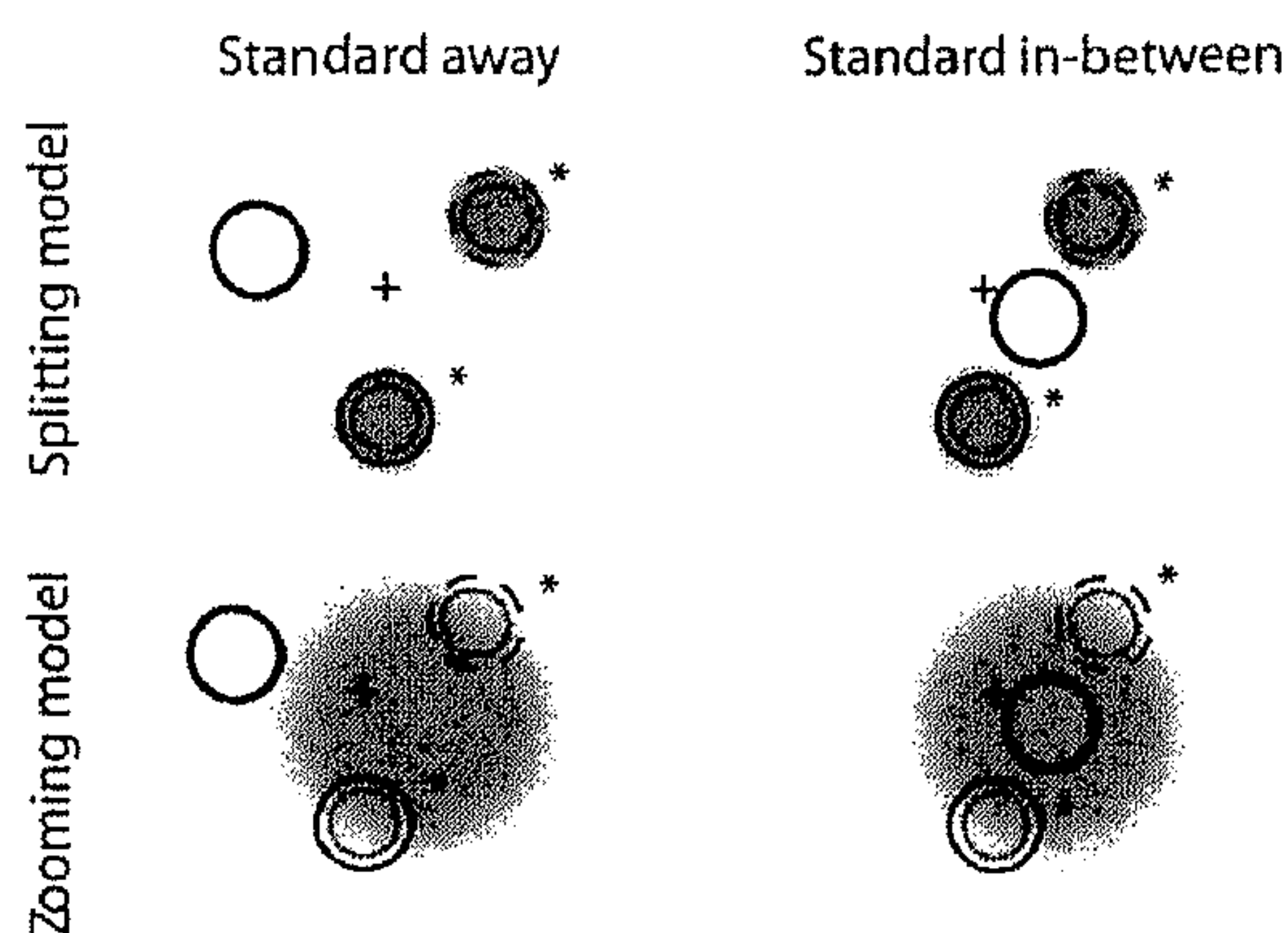


Figure 4. Experimental conditions used by Scharlau (2004) to test zoom-lens and splitting models. Gray scale gradients indicate attention distributions after cueing variable target and distracter according to splitting model (top row) and zooming model (bottom row). Not to scale, other conventions as in Figure 3.

All cueing conditions produced an equal rightward shift of the psychometric distribution and thus matched prior entry relative to the no cueing condition. Hence, in the standard in-between condition, there was no attentional enhancement of the standard target. Control experiments demonstrated that the color of the (masked) cue was irrelevant, and that the results could also be replicated with visible rather than masked cues. The results falsify the zooming theory, and indicate that under the conditions of the present experiment, attention can be allocated in parallel to non-contiguous locations, as predicted by the splitting model of attention.

Using distracter processing to study division of attention between two locations

In previous sections, the presence of attention in-between cued locations was probed by assessing the extent to which the processing of stimuli in-between those locations would be enhanced. Kramer and Hahn (1995) tested the possibility of non-contiguous division of the spotlight of attention by testing the extent to which truly interfering stimuli in-between cued locations would be ignored. To that aim, two target letters (cued by frame stimuli) were presented together with two distracter letters in-between, and the task of the subject was to indicate whether the target letters were matching or non-matching. The frames were presented mirror-symmetrically around the vertical meridian in the upper or lower quadrants of the

visual field, above or below fixation. The two target letters were presented inside the frames 150ms after onset of the frame cues, together with the two distracter letters in-between the target-locations. The display with letters remained present briefly (60ms) but they were not masked. Eye fixation was not monitored. The four letters were separated by about 2.5° , and had the same eccentricity (8.5°). The cue provided by the frame had an exogenous component (related to its onset 150ms before stimulus onset), and an endogenous component (related to the expectation of the target within the frame).

The four letters were either presented as sudden-onset stimuli (presented at once on a blank screen), or as non-onset stimuli. The non-onset stimuli were 'digital eights' (the number eight consisting of seven line segments) presented *together* with the frame cues in the locations of the target and distracter stimuli, and modified after 150ms into letters by removing segments of the digital eight. For example, the 'digital eight' became an 'E' by removing two segments. The non-onset stimuli were introduced to remove exogenous effects associated with sudden-onset stimuli.

Split focus and zoom lens models of attention make different predictions about the processing of the distracter letters in-between cued locations. According to the splitting model, the identity of distracter letters would not influence target-target matching responses, but according to the latter model it would. According to the zooming model, a subject's attention would not only be directed to the targets but also to the in-between distracters, and a distracter matching one of the targets would therefore facilitate a match response independently of the fact that the targets matched or not, and non-matching distracters would facilitate a non-matching response regardless of whether targets matched or not.

The reaction time and accuracy data for the sudden-onset conditions supported the zooming model. They showed that distracters matching one of the targets decreased the reaction time and increased accuracy, whereas non-matching distracters led to the opposite. Interestingly, results in the non-onset condition supported the splitting model of attention. In that condition, reaction time and accuracy did not depend on the matching or non-matching status of the distracters, and in fact, results remained the same when distracters were not presented at all. This indicates that in the non-onset condition, subjects could ignore the distracters and divide their attention non-contiguously between the two targets. By contrast, in the sudden-onset condition, the exogenous cues associated with the onset of all 4 stimuli may have overridden the attentional cues associated with the two frames presented earlier in time, leading to interference between targets and distracters. The strength of exogenous cues has been demonstrated in a number of studies (Theeuwes, 1995; Jonides & Yantis, 1988; Gottlieb et al. 1998; Kusunoki, et al. 2000; Donk & Theeuwes, 2001; Enns et al. 2001; Franconeri & Simons, 2005). In the non-onset condition, the small physical stimulus changes associated with the gradual modification of digital-eights into targets and distracters lacked the strength to override the frames' cueing effects.

The result in the non-onset condition (no interference from in-between distracters) is indicative of separate foci of attention, especially since masking the stimuli after 60ms strongly reduced the possibility of serial movements of a focus of attention. Nevertheless, to increase support for parallel processing in separate foci of attention, Hahn and Kramer (1998) extended the experiment of Kramer and Hahn (1995) by including a manipulation of stimulus quality, using otherwise similar displays and timing parameters. Displays were presented containing two target letters that were both of low quality (degraded), both of high quality or of low and high quality. A serial model predicts that total processing time would be the summation of the processing times for the individual stimuli, with the longest processing time for two low quality visual stimuli. A parallel model predicts that the total processing time equals the processing time of the stimulus that requires the longest processing time. In agreement with the previous experiment by Kramer and Hahn (1995), the results with sudden-onset stimuli supported the serial model, while results with the non-onset stimuli supported the parallel model. Thus, with sudden-onset stimuli, one of them would always capture attention first, leading to a re-allocation of attention to process the other target. Possibly, serial shifts of attention are triggered by stimuli with strong exogenous cues, and parallel processing in separate foci may be obtained in the absence of strong exogenous cues. The same results were obtained irrespective of the presence or absence of masking.

Taken together, the results suggest that under specific conditions, two non-contiguous locations can indeed be monitored in parallel by attentional mechanisms, but that attention deployment is very dependent upon stimuli, task requirements, and experimental design. The complexity of the determinants of attention deployment is underscored by a comparison of the above-discussed data with those from Scharlau (2004), whose experiments with *sudden* onset stimuli *did* appear to produce support for multiple foci of attention (see Section 2.2).

Beyond Serial Scanning, Zooming and Splitting: Flexible distributions of attention that include enhancement and suppression

Previous experiments used displays with fairly limited numbers of items. Awh and Pashler (2000) performed a number of experiments to test how attention could be divided among potential target locations in more complex displays with large numbers of distracters (Figure 8). Subjects fixated in the middle of a screen when presented with two cueing stimuli ('=' signs, presented for 750ms) that marked the positions where target numbers (two digits) would appear. Immediately following these two cues, a 5x5 array that contained 23 letters and 2 target numbers was presented, centered on the fixation spot (118ms, with some variability between subjects as exposure was adapted to performance). Eye fixation was not monitored. Only the 4 corners of the central 3x3 elements in the array could contain target digits, pairs of which could be cued. On 80% of the trials, the two positions where a digit could appear were cued validly, on 20% of the trials the

positions were cued invalidly. When the target positions were cued invalidly, one target digit appeared in a location between cued locations, and the other in a second position further away from cued locations. After the presentation of the array, all characters were masked (by # signs). After the mask, a cue was given to the subjects to indicate the two (cued or uncued) locations in the array for which a target-report was desired. The accuracy of the subjects' report was measured.

The accuracy data showed that on invalidly cued trials, accuracy of target reports for both invalidly cued positions decreased strongly compared to accuracy for validly cued positions. In particular, the strong reduction in accuracy of digit reports in locations in-between cued locations suggests that processing of the cued locations was enhanced, while the region in-between were ignored, in agreement with a splitting model of attention. In further experiments, similar results were obtained with non-verbal stimuli.

In their discussion of the data, Awh and Pashler (2000) proposed that the absence of attentional enhancement in-between cued positions could either reflect a mere absence of attentional resources at uncued positions, or an active suppression of irrelevant information. To test these hypotheses, the authors left out the 23 distracter letters, and only two target numbers were displayed after the cueing was done. Compared to the validly cued positions, there was only a slight decrease in accuracy in the un-cued, in-between positions, and although this decrease was still statistically significant, a large drop in accuracy was only seen in the far-away un-cued position. This suggests that the large drop in accuracy in the in-between un-cued positions in the previous experiments with distracters was to a large extent due to an active suppression mechanism, and not to the mere absence of attentional modulation at the un-cued positions. The contribution of distracter suppression to the shape of the attention distribution is in agreement with findings from Eriksen and Yeh (1985) (Section 2.1). These data indicate that in response to task requirements, an attentional distribution is set up that can include regions of enhancement as well as regions of suppression.

If the cueing of relevant positions has the effect of filtering out noise (distracters) in other positions, the mask itself could also be considered as noise. The authors therefore re-ran the previous experiment, leaving out all masking stimuli (# signs), except those following the target numbers. In this experiment, performance in the un-cued location in-between the cued target locations and in the two cued locations themselves was identical, and only a small, significant performance drop was found in the far-way un-cued position. Thus, leaving out the masking stimuli strengthened the trends in the data found after removing the distracters. This confirms that in displays with noise around the target locations (distracters or masking stimuli), the absence of attentional enhancement in non-cued locations does not simply reflect an absence of attention (and a division of attention between cued locations), but rather, an active suppression of irrelevant information. Thus, the deployment of attention over non-contiguous locations in space depends at least in part on the active suppression of noise in the visual field.

Therefore, an interplay of enhancement and suppression may determine the ultimate shape of the attention distribution. The existence of suppression outside the attended locus has been demonstrated by many psychophysical studies (Caputo & Guerra, 1998; Smith, Singh, & Greenlee, 2000; Müller & Hubner, 2002; Slotnick, Hopfinger, Klein, & Sutter, 2002; Slotnick, Schwarzbach, & Yantis, 2003; Müller & Kleinschmidt, 2004; Müller, Mollenhauer, Rösler, & Kleinschmidt, 2005), and may be related to negative BOLD in regions of the visual field from which attention is withheld (Smith et al. 2000; Shmuel, Yacoub, Pfeuffer, Van de Moortele, Adriany, Hu, Ugurbil, 2002; Smith, Williams & Singh, 2004).

The data from Awh and Pashler (2000) also show that the attentional distribution of suppression and enhancement is dynamically updated as stimuli change in time, as suggested by the effect of the mask on attentional performance. Hence, the cues that are used in many attentional paradigms to control attention are only one of many factors determining attentional allocation, and the distribution of attention can be updated, even within the course of a trial, to reflect changed task demands and the availability of new cues.

The above-described experiment from Awh and Pashler (2000) without distracters and masks demonstrated equally good target discrimination performance in the two cued positions and the in-between un-cued positions, and slightly but significantly reduced performance in far-away positions. Under the described conditions, these data indicate an enlarged region of attentional enhancement including the two cued locations and in-between locations, in agreement with a zoom lens model of attention. By contrast, data obtained with the prior entry paradigm from Scharlau (1999) suggested no difference in attentional processing between a standard in-between and a standard away condition, in agreement with a splitting model of attention, and falsifying a zooming model. Possibly, a target discrimination paradigm might be more sensitive to detect variations in the distribution of attention than a prior entry paradigm.

Are sequentially presented stimuli processed serially?

In the experiments discussed in previous sections, the targets were usually presented simultaneously. A comparison of sequential and simultaneous presentations, however, could be highly informative to determine the contribution of serial or parallel processes. Shiffrin and Gardner (1972) proposed that the existence of a single focus of attention (and thus the impossibility to select two non-contiguous locations simultaneously) would decrease performance if targets would be presented simultaneously in different locations (e.g., in a single 100ms display) compared to sequentially (e.g., two displays of 100ms). Testing this idea would require a simultaneous condition that is challenging, thus stretching attentional resources. If the low performance in a challenging simultaneous condition were due to insufficient time and resources to process the target with a single focus of attention, a sequential presentation would lead to doubling of the

time devoted to each target, and performance in the sequential condition would increase (serial model). If the low performance in a challenging simultaneous condition were due to insufficient time and resources to process the target with two foci of attention operating in parallel, a sequential presentation would not influence the time devoted to each target, and performance would remain unchanged (parallel model with fixed amount of resources in each focus).

Bichot et al. (1999) tested these ideas in an experiments in which subjects fixated a fixation cross in the middle of a screen (fixation was not monitored) while each trial presented a rapid succession of 8 displays (140ms / display) containing two digits. The location of the two digits was fixed in each display (one 2.5° to the left and one 2.5° to the right of fixation), creating two Rapid Serial Visual Presentation (RSVP) streams left and right from fixation. In each RSVP stream, red digits were targets and green digits distracters, or *vice versa* (randomized over subjects), and subjects had to report the highest target number presented in the course of the 8 displays. Emphasis was placed on accuracy; there was no time pressure. In the simultaneous condition, each of the 8 displays contained either two targets, or two distracters. In the sequential condition, a display showed a target on the left of fixation and a distracter on the right, or the opposite, and target and distracter were alternating on subsequent displays. Conditions were randomized, and within each condition, the series started randomly with one or the other of the two possible displays (see Figure 5).

Bichot et al. (1999) found that the error rates for the sequential condition were in the same range as the error rates in the simultaneous condition, supporting the idea of a truly split focus of attention. Importantly, the task proved to be challenging, as witnessed by an error rate of approximately 45% in the simultaneous condition. A lack of errors in the simultaneous condition would leave open the possibility that there was already sufficient time to process the two simultaneous targets with a serial mechanism, such that a lack of a performance increase in the sequential condition would be hardly surprising. Hence, if the 45% error rate in the simultaneous condition were due to a serial shifting mechanism under time pressure, the sequential presentation of the targets would have constituted an advantage. The lack of an increase of accuracy in the sequential condition, therefore, argues against a serial mechanism.

In principle, the results could also be accommodated by a zooming model of attention. An additional experiment, however, suggested at the most a limited decrease of performance when RSVP streams of distracters were shown at fixation in-between the two RSVP streams containing the target stimuli. Hence, the results taken together argue in favor of a relatively stable split of the focus of attention, each separate focus having too little resources [see also Bichot (1999) in Section 2.2, Figure 2C].

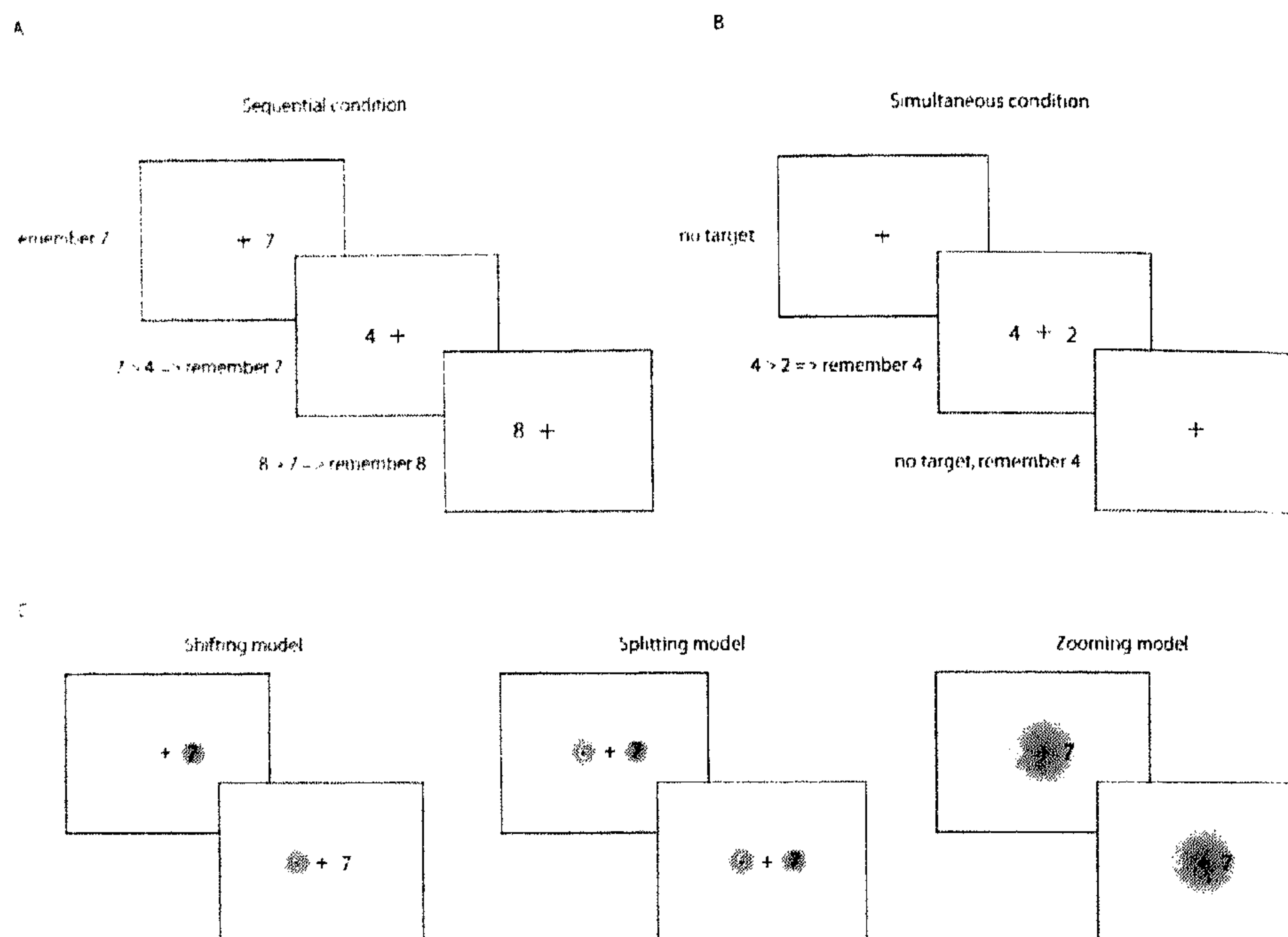


Figure 5. Overview of sequential and simultaneous conditions in experiments of Bichot et al. (1999), and summary of hypotheses derived from three attention models. A) Sequential condition, in which subjects had to report the highest target number that had occurred in a series of displays (three out of 8 displays shown). Targets (here illustrated in black) differed in color from non-targets (here illustrated in grey). Cognitive operations are indicated to the left of each display. B) Simultaneous condition, in which targets were paired within a display. Conventions as in (A), for further details see text. C) Possible attentional distributions predicted from three models of attention (splitting, shifting, and zooming). Attentional modulation is indicated by radial grey-level areas (with darker levels indicating stronger enhancement). Zooming lens models were tested by placing distracter stimuli at fixation and other locations in-between the target locations (not illustrated).

Interestingly, a trial-by-trial analysis in a control experiment also revealed some flexibility of the distribution of attention within trials. In this experiment, gratings were given target or distracter status based on their color. Of the two target gratings, presented simultaneously or sequentially in two frames, the target with an orientation closest to vertical had to be reported. With these stimuli, the same pattern of data was observed (no advantage for sequential presentations), arguing against serial attention shifts. Nevertheless, there was some evidence for flexibility in the distribution of attention. Trials in the experiment with grating stimuli consisted of only two frames, such that in the simultaneous condition a frame with two distracters predicted the presentation of two targets in the next frame. Subjects used that information, since simultaneous trials in which the first frame presented two distracters produced less errors (24%) than simultaneous trials in which the first frame presented the two targets (33%). Furthermore, in a condition in which

the targets were presented in the same location in successive frames (which was avoided in other experiments described in this section), only 17% errors were made. Apparently, when given the opportunity, subjects use information from preceding stimulus displays to update expectations on the probability of target presentation, and updated their attention distribution accordingly. This argues in favor of a parallel model of divided attention, with some flexibility to permit on-line fine tuning to accommodate changed processing needs during a trial or stimulus presentation.

Allocation of attention to object features

As evident in this review, a majority of studies have used spatial cues to test whether attention can be divided among multiple, non-contiguous locations (Eriksen & Yeh, 1985; Pan & Eriksen, 1993; Hahn & Kramer, 1998; McCormick, Klein & Johnston, 1998; Awh & Pashler, 2000; Scharlau, 2004). There is strong evidence, however, that the distribution of attention can be guided by surfaces, objects, and object-features (e.g., Duncan, 1984; Davis, Driver, Pavani, Shepard, 2000). For example, Duncan (1984) demonstrated that it is easier to switch between two features belonging to the same object than between two features belonging to separate objects, for equal spatial distances between the features. This shows that objects and surfaces can guide the distribution of attention, in addition to location cues. In addition, the effects of 'location' cues are often at least in part determined by their status as an object. For example, a frame cue delineates a surface, and has more object status than a small isolated dot or line segment. In line with that idea, attentional benefits resulting from cueing a location with a solid line segment spread far beyond the cue location (Rizzolatti, Riggio, Dascola, Umiltà, 1987; Umiltà, 1991), while attentional benefits from cueing with a frame are restricted to the regions encompassed by the frame (Hughes & Zimba, 1985, 1987). Thus, the distribution of attention at a location cued by a frame is determined both by spatial and object-related information. By the same token, directing attention to an object feature will result in a specific spatial distribution of attention. For example, when attention is directed to the red color of objects irrespective of their location, this will still result in a particular spatial distribution of attention over locations in the visual field on any given trial (Cave & Pashler, 1995; Bichot et al., 1999, 2005). In conclusion, in most studies we have discussed the spatial distribution of attention was guided by a mixture of spatial and object-related cues.

Electrophysiological studies

The psychophysical evidence reviewed in the previous section does suggest that under particular conditions, the visual system can divide attention into separate foci, each simultaneously affording a processing advantage to stimuli placed

within them, compared to stimuli placed outside. The aim of Woodman and Luck (1999), however, was to create conditions that favored serial attention shifts between a pair of targets, and to measure a physiological correlate of these shifts. As a correlate of attention, they used the *N2pc* ERP (Evoked Readiness Potential) component, which is a negative-going (N) component that reaches its minimum 200-400ms after stimulus onset (2), and is recorded posteriorly (*p*) on the scalp over the hemisphere contralateral (*c*) to the hemifield to which attention is directed. Modulations in the negativity of this component have been shown to be related to attentional selection and to changes in covert attention (Luck & Hillyard, 1994a, 1994b).

If subjects indeed used a shifting strategy in attending to two targets, it would be impossible to correlate changes in the *N2pc* component with changes in the location of focused attention unless one knew where the subjects' attention was directed first. Thus, Woodman and Luck (1999) devised a strategy to influence the order in which subjects would search multiple targets. Subjects fixated in the middle of a screen filled with outlines of squares, one side of which (left, right, up or down) showed a small gap. There were 21 black distracter squares spread in the visual field, and in addition, in each quadrant a differently colored (red, green, blue or violet) square was presented, which were potential targets. Eye fixation was not monitored. The subjects' task was to identify the target with a gap in the left rib. Before a trial began, subjects were informed which two colors were potential targets, and they were also informed that for one particular color the probability it would be a target was much higher (75%, T75) than for the other color (25%, T25). For example, if the red square was in the left hemifield and subjects knew there was a 75% chance it would be the target, compared to 25% for a green target on the right, subjects would be biased to direct their attention to the red stimulus first (left hemifield). If it would turn out that the red stimulus was not the target, attention would then be directed to the green stimulus (right hemifield). Under this hypothesis, a modulation in the negativity of the *N2pc* component would be expected, revealing a shift of attention. Alternatively, it is possible that knowledge about target probabilities would trigger a parallel strategy with a stable distribution of attention, in which more resources would be allocated to one location than to the other. Under this hypothesis, the negativity of the *N2pc* component associated with target stimuli in a particular location would be independent of target status (T25 or T75), because no change in allocation of attention would take place. Subjects were informed of the target probabilities, but were not instructed which strategy to use.

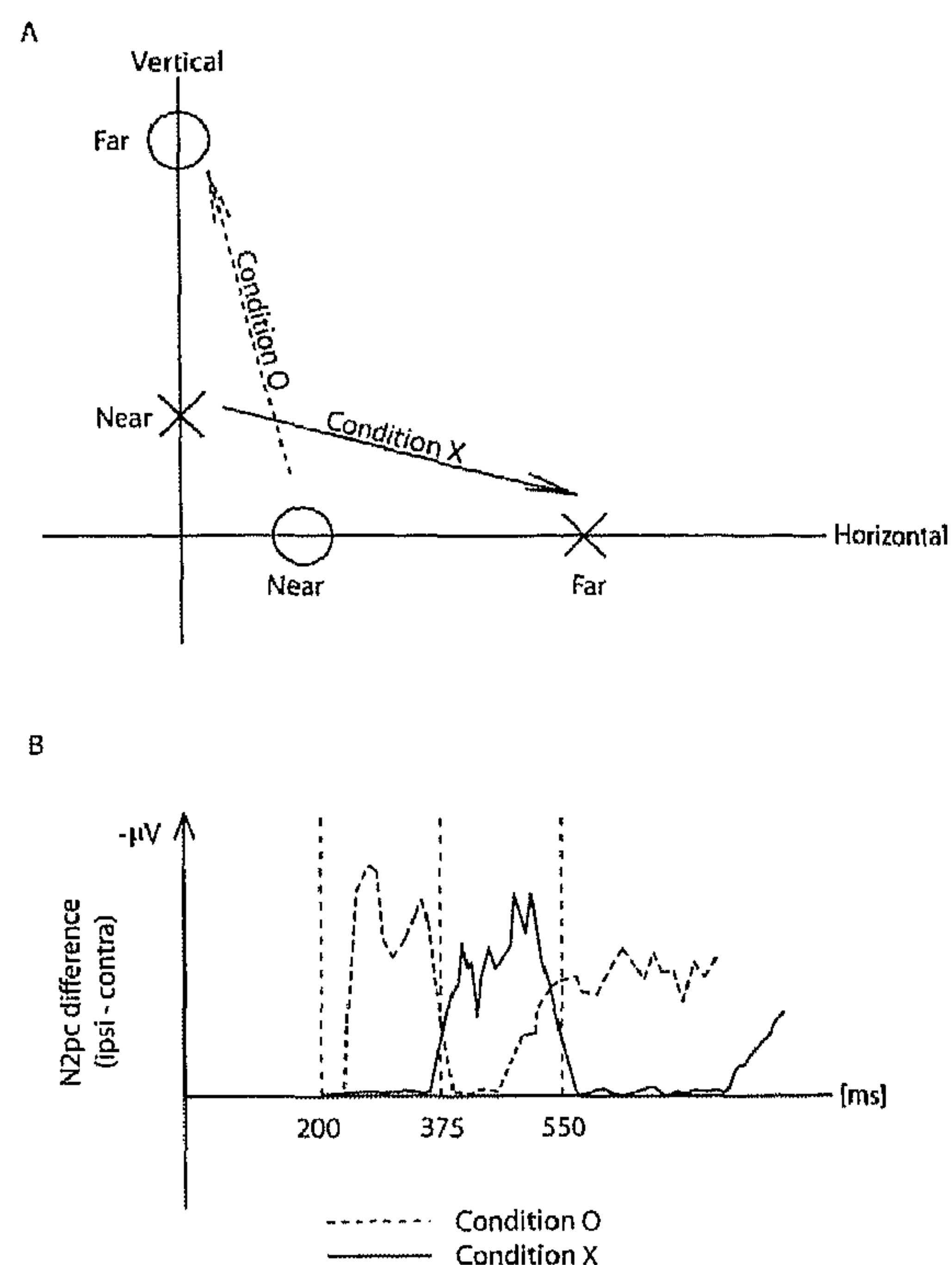
A comparison of the *N2pc* recorded from left and right halves of the scalp when T75 and T25 were placed in opposite hemifields showed that first (200-300ms) the *N2pc* contralateral to T75 was more negative than the *N2pc* contralateral to T25, while later on (300-400ms) this became reversed. This result is compatible with a movement of attentional resources from one location to

another one, and rejects the hypothesis of a fixed parallel distribution with unequal resources allocated to the two target locations.

To provide further evidence against stable parallel models, Woodman and Luck (2003) performed an elegant experiment with 4 pairs of potential targets (which were square outlines with potentially a gap in one of their ribs, as in the previous study). Each pair of targets had a specific color, and for each pair, one member of the pair was placed at a small eccentricity (near) on one principle meridian and the other at a large eccentricity (far) on the other principle meridian (Figure 6A). Color was used to direct attention to a pair of potential targets, and search order within the pair was biased by the difference in eccentricity (the candidate-target at lower eccentricity being more likely to be scrutinized first). The authors analyzed the results from the subset of target-absent trials in which none of the two members of the pair were a target, which were interspersed with target-present trials during data collection. While in the previous study (Woodman & Luck, 1999), the lateralization of N2pc components reflected relative allocation of attention to items in opposite hemifields, the N2pc components in the present design were un-influenced by attention to candidate targets on the vertical meridian. The N2pc components therefore only reflected 'absolute' attention to items on the horizontal meridian. Thus, after selection of a pair, more negativity of the N2pc component contralateral than ipsilateral to the horizontal meridian item during a time window the subject was paying attention to vertical meridian item would indicate parallel attention to both items in the pair. According to a serial model, no difference would be expected between N2pc components contra and ipsilateral from the item on the horizontal meridian during the time subjects inspect the vertical meridian item. For example, suppose a subject selected a near item on the vertical meridian and a far item on the horizontal meridian (Figure 6A). Despite a potentially asymmetrical distribution of attention corresponding to the bias to process the near vertical item before far horizontal item, a parallel model would nevertheless predict a correlate of the simultaneous selection of the items in the pair. This means that –from the moment the pair was selected- the amplitude of N2pc components measured contra-laterally from the horizontal meridian item would become more negative compared to the magnitude of the ipsilateral N2pc component. A serial model would predict a correlate of the bias to move a focus of attention from the near vertical item to the far horizontal item. In particular, a serial model would predict that during attention to the vertical near item, there would be no difference between the N2pc components measured contralaterally and ipsilaterally from the horizontal far item, and that only after moving attention to the horizontal far item, the N2pc component contralateral to the horizontal far item would become more negative than the ipsilateral N2pc component. The data provided support for the second alternative. The dashed line in Figure 6B shows that the bias to select a horizontal near item before a vertical far item was associated with the contralateral N2pc component being more negative than the ipsilateral component during the 200-375 interval of the ERP,

followed by a lack of differences between contra and ipsilateral components during the 375-550ms interval of the ERP. The solid line in Figure 6B shows that a bias to select a vertical near item before a horizontal far item was associated with an absence of a difference between contralateral and ipsilateral N2pc components during the 200-375ms interval of the ERP, followed by the contralateral N2pc becoming more negative than the ipsilateral component during the 375-550ms interval of the ERP.

Figure 6. Summary of main conditions and results of Woodman & Luck (2003). A) Principal experimental conditions. Condition O (circles): the two locations where targets could be presented were near the fixation point on the horizontal meridian, and far from fixation on the vertical meridian. Condition X (crosses): (potential) target locations were far from the fixation point on the horizontal meridian, and near fixation on the vertical meridian. On average, completion of processing of 'near' items should precede processing of 'far' items, which in a serial model would correspond to a shift of attention indicated by the arrows. B) Schematic illustration of obtained results. N2pc difference (N2pc re-corded ipsilateral versus contra-lateral to the horizontal-meridian target location) plotted as a function of time in condition O and X shown in (A). Additional target locations left and below fixation revealed similar results (see Woodman & Luck, 2003; not illustrated in figure). For further details, see text.



Taken together, the electrophysiological data suggest that movements of focused attention can be completed within possibly as fast as 100ms (Woodman & Luck, 1999), to about 150ms (Woodman & Luck, 2003). Woodman and Luck (1999, 2003) interpreted these findings as evidence for a serial model and against static inflexible parallel models. The authors acknowledged, nevertheless, that in principle any serial model can be emulated by a *flexible* parallel model, but they did not give this possibility serious consideration. However, there is no good *a priori* reason to exclude the possibility of flexible parallel operations, especially if this flexibility is biologically plausible. For a parallel model to emulate the serial-looking behavior shown by Woodman and Luck (1999, 2003), attentional reallocation would have to be fed back in parallel to retinotopic visual areas in about 100-150ms. This time window is compatible with other demonstrations of

parallel feedback influences in visual cortex (for review, see Lamme & Roelfsema, 2000). Direct evidence for much faster item-by-item switches in attention of a speed that would be incompatible with such feedback operations (on the order of 10-30ms) is entirely lacking. Therefore, it seems most parsimonious to interpret the data as support for a parallel model with a moderately flexible distribution of attention. In agreement with this idea, electrophysiological evidence in monkeys suggests that shifts in attention are associated with a simultaneous down-regulation of attention at one location and up-regulation of attention at another location (Khayat et al. 2006), and the same has been reported for shifts in attention between features of a single object (Motter, 1994). The up and down regulation of N2pc components in the studies of Woodman and Luck (1999, 2003) may reflect similar simultaneous operations. It is also worth noting that Bichot et al. (2005) did find a correlate of parallel attentional enhancement of a set of candidate targets in a primate recording study, and that the absence of such correlate in Woodman and Luck's (2003) study may be related to experimental design or limitations in sensitivity of the ERP technique.

Another strategy to investigate the allocation of attention in divided attention tasks is to assess attentional modulation in regions of the visual field in-between target locations. An ERP study from Heinze, Luck, Munte, Gos, Mangun and Hillyard (1994) found attentional enhancement of the processing of probe stimuli presented in-between relevant stimuli, which were interpreted as support for a zoom lens theory of attention, and which were also compatible with a moving focus of attention. It is possible, however, that these results primarily reflect limitations in spatial resolution of the ERP technique used.

Müller, Malinowski, Gruber and Hillyard (2003) reinvestigated the possibility of dividing attention by using the Steady State Visually Evoked Potential (SSVEP), which is a sustained ERP component elicited by a flickering stimulus in the visual field. SSVEPs oscillate at the frequency of the flickering stimulus, and therefore SSVEPs associated with different stimuli can be separated from each other if the different stimuli flicker at different frequencies. In this way, the low spatial resolution of the classical ERP technique can be circumvented, and signals from spatially nearby stimuli can be segregated. Previous studies had already shown that directing attention to a flickering stimulus increases the amplitude of the SSVEP (Morgan, Hansen, & Hillyard, 1996; Müller et al. 1998; Müller & Hubner, 2002). This suggested that directing attention to one of several stimuli flickering at different rates should be revealed by an increase in the amplitude of the SSVEP coupled with the attended stimulus.

In Müller et al.'s (2003) experiment, four small flickering squares were presented on the horizontal meridian; two to the left of fixation (at approximately 5° and 10° of eccentricity), and two placed symmetrically to the right. Subjects were instructed verbally to attend to two of the squares. In each square, a randomized Rapid Serial Visual Presentation (RSVP) sequence of symbols was presented at a rate of 181ms per display. The subjects' task was to attend to two

squares and signal the co-occurrence of the same symbols within the squares; a task that was designed to direct equal attention to the two squares. The 181ms display duration is considered to strongly decrease the possibility of attention shifts during display duration (see Introduction). Attention directed to two spatially segregated target stimuli resulted in increased amplitudes of the SSVEPs coupled to these stimuli compared to the SSVEPs corresponding to distracter stimuli positioned in-between the target stimuli. These findings suggested that attention could be divided between two separate locations in a parallel manner (splitting model), and are not easy to reconcile with serial shift models and zoom lens models of attention.

Whether or not there was simply a lack of attentional enhancement of information or active suppression in-between relevant regions was not addressed in Müller et al.'s (2003) study. Hopf, Boehler, Luck, Tsotos, Heinze and Schoenfeld (2006) used MEG and found a narrow region of suppression around a region of attentional enhancement in a pop-out task. Electrophysiological demonstrations of suppression in divided attention tasks that require active suppression of distracters in-between enhanced regions (Awh & Pashler, 2000; Section 2.5) are currently lacking.

Altogether, electrophysiological experiments indicate that attention can be divided in a parallel manner, and that changes in attentional allocation are accomplished by simultaneous up-regulation and down-regulation of attention in different locations as required by the task. The data do not provide a compelling reason to hypothesize the existence of a specific serial mechanism, in which a single focus of attention moves between relevant locations.

fMRI studies

Psychophysical studies suggest that subjects will shape their attention distribution into two peaks if necessary for the task, but that relatively quick modifications of the attention distribution will be carried out even within trials to accommodate changed processing demands. The electrophysiological studies reviewed above are entirely compatible with that conclusion.

fMRI studies have proven to be an effective tool to study attentional enhancement in retinotopic visual areas, which can be demonstrated in the presence (Tootell et al. 1998; Brefczynski & DeYoe, 1999; Müller, Bartelt, Donner, Villringer, & Brandt, 2003) or in the absence of attended stimuli (Kastner, Pinsk, De Weerd, Desimone & Ungerleider, 1999). Furthermore, studies by Somers, Dale, Seiffert, and Tootell (1999) and by Müller and Hübner (2002) have shown that the shape of a region selected by attention can be complex. Because of the relatively good spatial resolution of fMRI, a number of fMRI studies have aimed to investigate attentional modulation in-between multiple regions of

attentionally enhanced processing, the presence of which is instrumental in distinguishing between serial shifting, zooming, and splitting models of attention.

McMains and Somers (2004) designed a fMRI experiment, in which subjects simultaneously monitored two of five available Rapid Serial Visual Presentation (RSVP) streams of small numbers and letters (one stream in each quadrant at 3.6° of eccentricity, and one in the middle of the screen at fixation; eye fixation was not monitored.). The RSVP streams that had to be monitored by the subjects were in two quadrants of different hemifields. In the RSVP streams, a new number or letter was presented every 173 ms. Subjects had to report when matching numbers were present simultaneously in the two attended RSVP streams. The other three RSVP streams (in the remaining two quadrants and in the middle of the screen) served as distracters. The fMRI data showed attentional modulation for the RSVP streams to which subjects had to attend for the matching, and no attentional modulation for the RSVP stream in between.

In a follow-up study, McMains and Somers (2005) used fMRI to find out whether there are extra costs associated with the deployment of attention over multiple locations compared to a unitary focus of attention, and whether splitting of attention is more economical compared to using a single large zoom lens of attention encompassing not only relevant locations but also distracter locations. Subjects were presented with five RSVP streams, one in each quadrant and one in the middle. Subjects fixated the middle RSVP stream, and covertly directed their attention (1) to a single peripheral RSVP stream (Figure 7A); (2) to two adjacent RSVP streams (Figure 7B); (3) to three adjacent RSVP streams (Figure 7C); (4) to two segregated RSVP streams (Figure 7D); (5) to a RSVP stream in a location not used in any of the 4 conditions above. The latter condition was used to measure an unattended baseline for the first 4 conditions. In all 5 conditions, subjects were required to monitor the attended RSVP streams for the appearance of a single target letter. The subjects' letter detection performance decreased with the number of streams that needed to be monitored. However, the performance when attending to two adjacent RSVP streams was similar to the performance obtained when two segregated RSVP streams were monitored. The magnitude of the fMRI attentional enhancement signal (reflecting a subtraction of attend-ignore conditions) closely mimicked the performance data. The more RSVP regions had to be attended, the lower the attentional enhancement in each region, irrespective of whether those regions were adjacent or not (Figure 7). Hence, for a constant number of locations (compare Figures 7B,D), the percent signal change was the same irrespective of whether those regions were adjacent or not. This indicates that the monitoring of multiple locations took place at the cost of more limited attentional resources in each location occupied by an attended RSVP stream. Furthermore, there was no enhancement of the fMRI signal in cortex corresponding to a non-relevant RSVP-stream, placed in-between two relevant RSVP streams (Figure 7D).

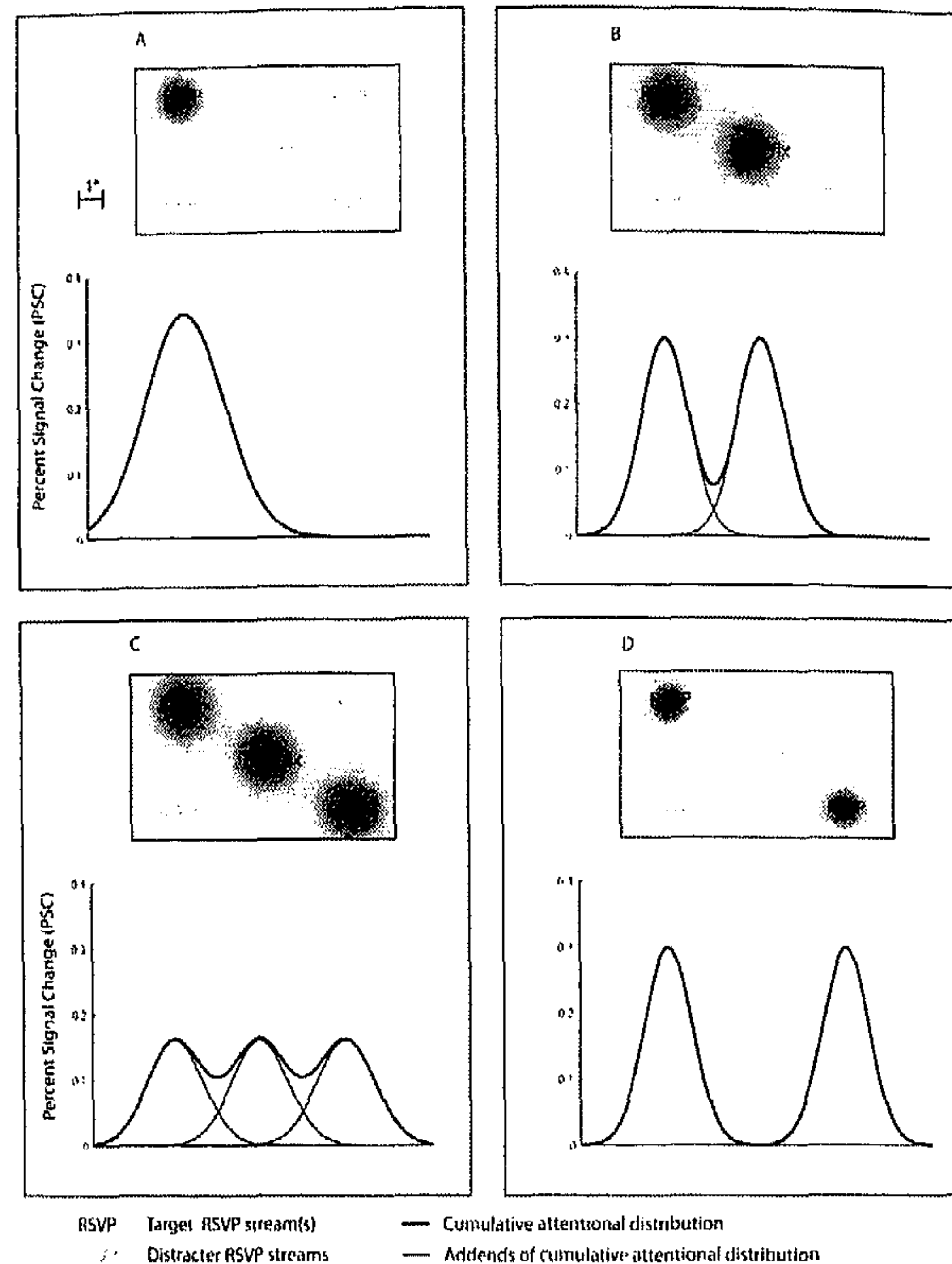


Figure 7. Principal experimental conditions and attentional distributions matching the obtained fMRI result in McMains and Somers (2005). (A-D) Four principal attention conditions with hypothesized distributions of attention. Each stimulus display contained five RSVP streams, and item-size was scaled for cortical magnification (drawings not to scale). Attention could be directed to a single target RSVP stream (A), to two adjacent RSVP streams (B), to three adjacent RSVP streams (C), or to two non-adjacent RSVP streams, separated by a RSVP stream in the middle (at fixation) which had to be ignored (D). The gray ellipsoids in the displays show theoretical distributions of attentional enhancement compatible with observed behavioral and fMRI evidence for enhancement compared to a passive-viewing baseline. Darker levels correspond to larger enhancement. Below the illustration of the stimulus display in each attention condition, a unidimensional cross-section is shown through the attention distributions shown in the stimulus display (A-D). Thin solid lines show the addends that after summation give the overall attention distribution shown in thick solid line. The peaks of the addends distributions and of the overall attention distribution match the observed attentional enhancement measured with fMRI. The width of the addends distributions was assumed to be identical in conditions shown in B-D, but in order to keep the total area under the overall distribution (fat line) constant in all conditions, a wider attention distribution was required in the condition shown in A (where all attention is directed to a single RSVP). The total surface under the overall attention distribution was kept constant, following the assumption of constant, limited resources of attention. The figure shows a decline of peak enhancement in regions of interest corresponding to the RSVP streams as more RSVP streams had to be attended. Panel D corresponds to the finding that fMRI can reveal separate attentional foci if RSVP streams are sufficiently far apart (7.2°) and if the middle RSVP stream has to be ignored. See text for further details.

The data are most easily understood in the context of a splitting (multiple-foci) model of attention as illustrated in Figure 7. In this model, separate Gaussian regions of enhancement (addends distributions in Figure 7 shown in thin solid line) can be assumed in the location of each attended RSVP stream. The peaks of

the addends distributions were chosen to match the attentional enhancement measured with fMRI (the fMRI data showed decreased enhancement as a function of an increase in the number of attended RSVPs). The widths of the addends distributions were chosen to keep the area under the overall, summated attention distributions constant (summed distributions in Figure 7 shown in thick solid line). This led to constant widths of addends distributions when two or three RSVP streams were attended (Figure 7B-D), but a much wider distribution when a single RSVP stream was attended (Figure 7A). According to this model, splitting attention to accommodate two instead of one target is done by reshaping a single, relatively wide distribution into two narrower distributions with only marginally smaller peaks. Splitting attention to accommodate three rather than two targets is done by a large reduction in the peak of the attention distribution in each relevant location without further changes in width. What the actual width and shape of the attention distributions was in McMains and Somers (2005) is unknown, because they only reported the amplitude of attentional enhancement averaged within predefined regions-of-interest corresponding to the location of the RSVP streams. Psychophysical studies suggest that the way in which attention is divided differs between paradigms. The psychophysical data from McMains and Somers (2005) indicate that the very small (but statistically significant) decline of (maximum) attentional enhancement when two rather than a single RSVP stream had to be attended is sufficient to induce a clear decline in attentional performance. This is in agreement with Castiello and Umiltà (1992) and McCormick et al. (1998), who used a frame-cueing paradigm, and found that directing attention to two rather than one location resulted in a significant decrease in performance (Section 2.2). Scharlau (2004), by contrast, who used a prior entry paradigm, concluded that a division of attention between two locations does not result in lesser performance in the two locations compared to a condition in which attention was directed to a single location (Section 2.3). More fine-grained analysis of the distribution of attentional enhancement with fMRI may be useful to test assumptions of divided attention theories in different paradigms.

At least three factors might contribute to the observation of separate distributions of attention in an fMRI study in which attention is directed to more than a single location: (1) the resolution of the attentional system itself, (2) the spatial resolution of the fMRI signal, and (3) the specific paradigm used. In at least one condition in McMains and Somers (2005), the resolution requirements were fulfilled and the paradigm used permitted the observation of two separate regions of attentional enhancement (Figure 7D). Active suppression of information in-between attended locations may have been important in achieving two separated attention distributions [see Awh and Pashler (2005), Section 2.5].

The lack of enhancement for a region in-between attended locations is incompatible with a zoom lens model of attention (Figure 7D). Likewise, the finding of equal attentional enhancement for two RSVP regions in adjacent locations (Figure 7B) and in non-adjacent locations (Figure 7D) is difficult to

reconcile with a simple zoom lens model of attention. Especially in the latter condition (Figure 7D) the distribution of attention would have to be extremely broad to accommodate the large attentional enhancement observed in far separated regions in the display. In general, zoom lens models would imply a tremendous waste of attentional resources because a continuous (Gaussian) distribution that could explain the attentional enhancement in separate locations would necessarily lead to attentional enhancement in large surface areas irrelevant for the task.

The fMRI evidence thus strengthens support for a multi-focus, parallel model of divided attention. The various features of the fMRI data point to a parallel modulation and division of attentional resources across the visual field, and would be difficult to explain with a serial model of attention characterized by a single, moving focus of attention.

Simulations

The reviewed data indicate that attentional distributions reflect task demands in both a spatial and a temporal manner. The distribution can involve several regions of enhancement in relevant locations, while regions in-between relevant locations show a lack of enhancement, or suppression. When task demands are stable over time, the distribution may remain stable during task execution. When task demands change during a trial, or when trials are set-up to make a re-allocation of attentional resources beneficial, then relatively fast changes in the distribution can occur. These re-allocations can occur without affecting the distribution in-between relevant locations, thereby excluding zooming models and rendering serial shifts of a focus of attention improbable (especially if stimulus exposure was limited). In addition, the reduction of attention available to relevant regions with increasing numbers of target regions takes place without affecting attention resources in-between relevant regions. Therefore, the distribution of attention in essence is parallel, though it is flexible enough to mimic an attentional behavior that looks serial. The opposite conclusion, namely that attention is allocated serially by moving a focus of attention, and that parallel behavior is mimicked by fast serial scanning of relevant locations or zooming is more difficult to reconcile with the data. In particular, the absence of attentional enhancement in-between multiple relevant locations argues both against a zooming lens model and quick serial shifting of a classical attention focus.

The idea of a flexible attentional distribution that would affect sensory processing in parallel throughout the visual field is in agreement with the notion of a saliency map, in which relevant regions in the (visual) environment are indexed in various spatial coordinate systems. This notion is supported by many neurophysiological studies of parietal cortex in monkeys (Gottlieb et al. 1998; Kusunoki et al. 2000; for review, see Colby & Goldberg, 1999; Andersen & Buneo, 2002), and used in many modeling studies of attention (e.g., Koch &

Ullman, 1985; Itti & Koch, 2000; Deco et al. 2002; Standage, Trappenberg, Klein, 2005a,b). The basic idea in these modeling studies is that the saliency map integrates bottom-up stimulus features with top-down attentional control (see Figure 4). The bottom-up input to the saliency map is retinotopically organized, which permits top-down input to modulate bottom-up input in parallel in multiple places in the map. In the above-referred studies, bottom-up inputs (stimuli) compete for processing resources within the saliency map, and the top-down signal (task-related instructions or strategies) biases the competition in favor of stimuli or locations that are relevant in the context of the behavioral task. The result of this bottom-up / top-down interaction is a distribution of salient locations in the saliency map. This distribution of saliency (attention) is then fed-back in parallel to sensory maps to resolve sensory competition (Desimone & Duncan, 1995). Such a model can emulate a single focus of attention, but it can also produce multiple foci of attention, with no attention allocated in-between, or with suppression in-between. Shifts of attention are accomplished by simultaneously upregulating modulation at one location, and down-regulating at another location, without affecting locations in-between. Capacity limitations in the model result in a reduction of modulation per total attended area when the number of relevant regions is increased (Figure 8).

Using a model inspired by Itti and Koch (2000, 2001) and Deco et al. (2002), Standage et al. (2005b) were able to reproduce the division of attention that was reported by Awh and Pashler (2000). The latter authors found that attention could be divided non-contiguously between target letters in a 5x5 array (see Figure 4), and the model of Standage et al. (2005b) successfully re-produced separate peaks in the attention distribution. Awh and Pashler (2000) also reported that excluding distracter letters surrounding the targets from the array reduced subjects' ability to divide attention discretely over the target locations (attention spilled over into in-between regions), but this result was not reproduced by the model: a discrete division of attention was obtained even when distracters were removed. Awh and Pashler (2000) have argued that separate peaks in the attention distribution only appear if noise must be suppressed in-between relevant locations. Likely, a number of factors influence the way in which attention is divided, including the distance between relevant locations, the extent to which the task is demanding (leading to enhancement), and the extent to which there is a requirement to ignore noise (leading to suppression). The interplay between all those factors must be included in the model, and it is not clear whether Standage et al. (2005b) have given sufficient weight to suppressive modulations. Furthermore, Standage et al.'s (2005a) model did not reproduce separate peaks in the distribution of attention when presented with the RSVP stimuli used by Müller et al. (2003). Hence, the model worked better for transient than for sustained attentional tasks, which suggests that careful consideration of temporal factors that govern connectivity will be required to expand the model to attention monitoring tasks with RSVP stimuli. Nevertheless, modeling efforts such as those by

Standage and colleagues support the principle of the saliency map as a substrate for flexible parallel distributions of attention. Conditions in which a model fails to mimic human performance can provide insight in specific network properties that must be present to capture the finer details of human attentional performance.

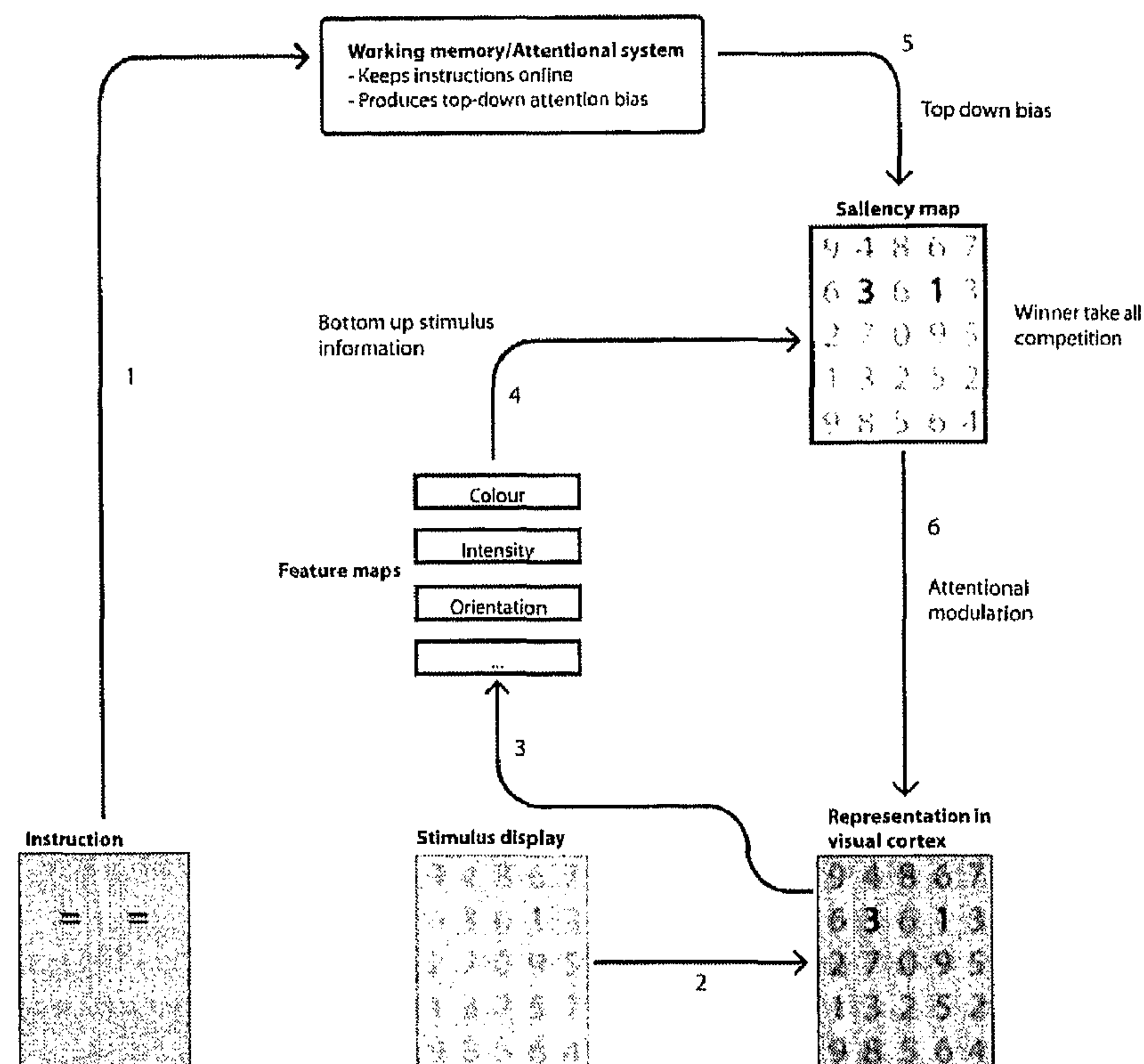


Figure 8. Application of a Winner Take All (WTA) attentional network to the study of Awh and Pashler (2000). The sequence of information processing is indicated by numbered arrows. (1) Subjects receive an instruction where to attend while fixating in the middle of the screen. This instruction is stored in working memory. (2) After the instruction, a 5 by 5 array of numbers is presented, and this display is represented in the visual cortex. (3) The visual information is processed in several feature maps. (4) The combined bottom-up information from the feature maps competes in a WTA fashion in the saliency map (posterior parietal cortex). (5) Working memory provides a top-down signal that biases the competition in the saliency map, leading to increased attentional enhancement in the locations of the relevant numbers. (6) The resulting attentional distribution in the saliency maps is then fed back to sensory areas, which leads to the resolving of sensory competition, and enhanced sensory processing of numbers in relevant locations. Consecutive iterations of this loop will contain the results of attentional enhancement achieved during past iterations, leading to further attentional enhancement in relevant locations, until an asymptote is reached. Attentional enhancement is symbolized by darkening of the digits.

Discussion

A number of articles reviewed above showed that it is possible to process visual information presented in two or more locations in parallel. In their ERP study, Müller et al. (2003) used SSVEPs and found evidence in favor of parallel processing. McMains and Somers (2004) did an fMRI experiment that showed attentional modulation in non-contiguous regions in visual space, indicating that parallel processing was possible. In addition, fMRI data from McMains and Somers (2005) demonstrated that capacity limitations revealed by increasing the number of target locations modulated the attentional distribution in the target locations only, without modulating regions in-between target locations, further supporting parallel processing. The combined psychophysical data from Castiello and Umiltà (1992), Kramer and Hahn (1995), Hahn and Kramer (1998), Bichot et al. (1999), Awh and Pashler (2000) and Scharlau (2004) also support the idea that subjects can use a parallel strategy. In all of these studies, there were no instructions or cues that would directly induce a serial strategy, and it appears that in the absence of such cues or instructions the default processing strategy is parallel.

Studies in which the authors interpreted their data in support of serial processing of information in multiple locations used cues in which a shift of attention was unavoidable because of an initial exogenous capture of attention in one location, because of instructions to expect a target more frequently in one location than in another, or because of other biasing procedures. In a psychophysical study of Eriksen and Yeh (1985), the use of a single exogenous cue to direct attention to two locations produced a shift of attention when the target was not found in the location most adjacent to the cue. Similarly, in the electrophysiological studies from Woodman and Luck (1999, 2003) an endogenous cue (instruction) or asymmetries in eccentricity strongly encouraged subjects to shift attention. Although Woodman and Luck (1999, 2003) considered their data as support for a serial shifting model, their data are also compatible with a moderately flexible parallel model (Bundesen, 1990). In particular, the data can be interpreted by assuming that a particular distribution of attention reflecting task demands at the beginning of a trial can be adapted within the course of a trial to accommodate changing processing needs. A series of studies from LaBerge and colleagues (LaBerge & Brown, 1986, 1989; LaBerge, Carlson, Williams & Bunney, 1997) and a study from Pratt and Quilty (2002) strongly support this interpretation. These studies used series of stimulus displays (exposure durations of 150-200ms, with short inter-stimulus intervals of 50ms) in which letters were packed side-by-side in dense horizontal arrays, with a letter in a particular position being a target. According to the ideas of LaBerge and colleagues, the presentation of a long horizontal array of place-holders (meaningless characters), followed by a long horizontal array of letters with a target in the middle, establishes a broad, symmetrical distribution of attention. This attention distribution is then disturbed

by presenting a smaller array with a target or a meaningless exogenous cue in a horizontal position overlapping with the long array. The detection performance of targets presented in subsequent small probe arrays presented in various horizontal positions reveals broad, asymmetrical distributions of attention that are incompatible with a moving attentional focus.

Some studies have shown attentional enhancement in-between attended locations, which could be taken as support for serial models of divided attention (zooming or scanning). However, divided attention tasks in which the cue or the instruction reliably predicted the location of the target (Castiello & Umiltà, 1992; Kramer & Hahn, 1995; Hahn & Kramer, 1998) always revealed non-contiguous regions of attentional enhancement. Only when cues or instructions did not reliably predict target locations, broad regions of attentional enhancement encompassing the targets have been reported (Eriksen & Yeh, 1985; McCormick et al. 1998; see also Posner et al. 1980). In the context sketched in the review, these latter results do not support serial models of attention (zooming or serial scanning). Instead, the combined results illustrate the way in which a parallel distribution can be set-up flexibly to maximize behavioral performance given particular task requirements and target probabilities.

In order to exclude a serial model as a possible interpretation of performance in divided attention tasks, many experimenters have restricted processing time to below the lowest limit at which one could still expect serial attention shifts between relevant locations. This critical limit has been derived from research with a variety of paradigms in humans (Krose & Julesz, 1989; Duncan et al. 1994; Ward et al. 1996; Woodman & Luck, 1999, 2003; Theeuwes, Godijn, & Pratt, 2004; for reviews see Moore et al. 1996; Egeth & Yantis, 1997) and monkeys (Motter, 1994; Khayat et al., 2006), and depending on the paradigm has been estimated to range from about 100 to 200ms, or sometimes more (see Introduction). The variability in this estimate shows how difficult it is to determine a precise processing time below which serial shifts would be excluded with certainty. However, limiting processing time as much as possible, and choosing a sufficiently high difficulty level of the task, are elements which in combination with other converging evidence can help to significantly reduce the possibility of serial shifts of attention in particular experiments. For example, the finding of separate regions of attentional enhancement (based on psychophysical and fMRI evidence), in a demanding divided attention task with short display durations (less than 200ms) argues against serial scanning by a focus of attention (McMains & Somers, 2004, 2005). In agreement with behavioral data, electrophysiological studies in monkeys (Motter, 1994; Khayat et al., 2006) and human subjects (Woodman & Luck, 1999, 2003) suggest that re-allocations of attention require 100-200ms.

Some behavioral studies could be interpreted as evidence for fast serial movements of attention (shifts from item-to-item in a few tens of ms). Specifically, studies of visual search showing increases of target response times

with the number of elements in conjunction search displays (20-30ms per item) might be used as evidence for a rapidly moving focus of attention (e.g., Treisman, 1982; see also Treisman & Gelade, 1980). However, parallel strategies can be used in conjunction displays to group items according to their features, and individual scrutiny of elements by covert attention is not required in these displays (Mordkoff, Yantis & Egeth, 1990; Treisman & Sato, 1990; Wolfe, 1994). Search arrays in which items must be scrutinized individually by covert attention show an increase in search time of about 100ms/item (Bricolo et al. 2002). In a divided attention task in Eriksen and Yeh (1985), subjects had to move their attention in a subset of trial from a cued location where the target had not appeared to a second location implicitly indicated by the same cue. Response times to the target in the implicitly cued location were delayed by only 50-80ms. When attention had to be moved to a non-cued location, responses were delayed by about 115ms. This might suggest that in experiments where targets can be expected in two locations with a high level of confidence, fast shifts in focused attention are possible [in Eriksen and Yeh (1985): 50-80ms]. However, in other divided attention tasks in which the cue or the instruction reliably predicted the location of the target (Castiello & Umiltà, 1992; Kramer & Hahn, 1995; Hahn & Kramer, 1998) attentional enhancement was observed in *non-contiguous* regions. To explain this result with serial shifts of focal attention would require that the focus of attention in a time span of just a few 10s of ms would be turned on, off, and back on again. We believe this is unlikely (see below).

Given the absence of direct physiological or compelling behavioral evidence for movements of a focus of attention in much less than 100ms, it is biologically plausible that attentional modulation of visual processing takes place by a parallel mechanism that can emulate serially-looking behavior through the action of feedback (Section 5, Figure 8). It is interesting to speculate about the essential difference in biological implementation between a mechanism that achieves sensory modulation through a serially operating focus of attention and a mechanism that achieves sensory modulation through parallel feedback. A neuronal network implementation of a rapidly moving focus of attention requires a minimization of the role of feedback, and might involve a mechanism in which a focus is initiated by a brief feedback interaction between a control network and retinotopic visual areas, after which the focus would move automatically, randomly, and in a continuous fashion through retinotopic maps, by virtue of lateral anatomical connectivity. In such a pure, serial mechanism, the attention focus therefore has to remain on (see Introduction). It is conceivable that such a mechanism would permit rapid scrutiny of individual elements (on the order of a few tens of ms per item). An interesting but controversial study by Horowitz and Wolfe (1998) (see Kristjansson, 2000; for review, see Kane, Poole, Tuholski & Engle, 2006) suggested that attention indeed moves randomly from item to item in search arrays (not keeping track of rejected distracters). However, it is difficult to see how a focus of attention would survive in the absence of a feedback

mechanism, and once one accepts a feedback mechanism it is difficult to see how the fundamental model of attentional modulation can be anything else but parallel. We propose that a focus of attention is most straightforwardly implemented by using the feedforward/feedback interactions available by default in the parallel architecture of the visual system (Lamme & Roelfsema, 2000). We therefore suggest that the limits on the speed of attention shifts are given by the architecture of these feedforward/feedback interactions, implying they can never be as fast as the 20-30ms suggested by some initial psychophysical studies of serial search. This means that terms such as 'serial scanning', and 'a focus of attention' (whether zoomed or split) can be used as a phenomenological characterization of data obtained in a subset of divided attention tasks, but that it is most parsimonious to think of the underlying mechanism as parallel. *'To shift or to divide?'*: it's just a rhetorical question.

References

- Andersen, R. A., & Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. *Annual Reviews Neuroscience*, 25, 189-220.
- Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal Experimental Psychology Human Perception and Performance*, 26(2), 834-846.
- Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature*, 303(5919), 696-698.
- Bichot, N. P., Cave, K. R., & Pashler, H. (1999). Visual selection mediated by location: feature-based selection of noncontiguous locations. *Perception & Psychophysics*, 61(3), 403-423.
- Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. *Science*, 308(5721), 529-534.
- Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nature Neuroscience*, 2(4), 370-374.
- Breitmeyer, B. G., & Kersey, M. (1981). Backward masking by pattern stimulus offset. *Journal Experimental Psychology Human Perception and Performance*, 7(5), 972-977.
- Briand, K. A., & Klein, R. M. (1987). Is Posner's "beam" the same as Treisman's "glue"? On the relation between visual orienting and feature integration theory. *Journal Experimental Psychology Human Perception and Performance*, 13(2), 228-241.
- Bricolo, E., Ganesini, T., Fanini, A., Bundesen, C., & Chelazzi, L. (2002). Serial attention mechanisms in visual search: a direct behavioral demonstration. *Journal of Cognitive Neuroscience*, 14(7), 980-993.
- Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, 47(3), 191-196.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, 97(4), 523-547.
- Caputo, G., & Guerra, S. (1998). Attentional selection by distractor suppression. *Vision Research*, 38(5), 669-689.
- Castiello, U., & Umiltà, C. (1992). Splitting focal attention. *Journal Experimental Psychology Human Perception and Performance*, 18(3), 837-848.
- Cave, K. R., & Pashler, H. (1995). Visual selection mediated by location: selecting successive visual objects. *Perception & Psychophysics*, 57(4), 421-432.
- Chastain, G. (1992). Time-course of sensitivity changes as attention shifts to an unpredictable location. *Journal of General Psychology*, 119(2), 105-111.

- Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual Reviews Neuroscience*, 22, 319-349.
- Davis, G., Driver, J., Pavani, F., & Shepherd, A. (2000). Reappraising the apparent costs of attending to two separate visual objects. *Vision Research*, 40(10-12), 1323-1332.
- Deco, G., Pollatos, O., & Zihl, J. (2002). The time course of selective visual attention: theory and experiments. *Vision Research*, 42(27), 2925-2945.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Reviews Neuroscience*, 18, 193-222.
- Deutsch, J. A., & Deutsch, D. (1963). Some theoretical considerations. *Psychological Review*, 70, 80-90.
- Donk, M., & Theeuwes, J. (2001). Visual marking beside the mark: prioritizing selection by abrupt onsets. *Perception & Psychophysics*, 63(5), 891-900.
- Driver, J., & Baylis, G. C. (1989). Movement and visual attention: the spotlight metaphor breaks down. *Journal of Experimental Psychology Human Perception and Performance*, 15(3), 448-456.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology*, 113, 501-517.
- Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, 369(6478), 313-315.
- Egeth, H. E., & Yantis, S. (1997). Visual attention: control, representation, and time course. *Annual Reviews Psychology*, 48, 269-297.
- Egely, R., & Homa, D. (1984). Sensitization of the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 778-793.
- Enns, J. T., Austen, E. L., Lollo, V. D., Rauschenberger, R., & Yantis, S. (2001). New objects dominate luminance transients in setting attentional priority. *Journal of Experimental Psychology Human Perception and Performance*, 27(6), 1287-1302.
- Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, 4(9), 345-352.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 433-449.
- Eriksen, C. W., & Murphy, T. D. (1987). Movement of attentional focus across the visual field: a critical look at the evidence. *Perception & Psychophysics*, 42(3), 299-305.
- Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: some results and methodological considerations. *Journal of Experimental Psychology*, 79(2), 1-16.
- Eriksen, C. W., & St James, J. D. (1986). Visual attention within and around the field of focal attention: a zoom lens model. *Perception & Psychophysics*, 40(4), 225-240.
- Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal Experimental Psychology Human Perception and Performance*, 11(5), 583-597.
- Estes, W. K., & Wessel, D. L. (1966). Reaction time in relation to display size and correctness of response in forced choice signal detection. *Perception & Psychophysics*, 1, 369-373.
- Franconeri, S. L., & Simons, D. J. (2005). The dynamic events that capture visual attention: A reply to Abrams and Christ (2005). *Perception & Psychophysics*, 67(6), 962-966.
- Frey, R. D. (1990). Selective Attention, Event Perception and the Criterion of Acceptability Principle: Evidence Supporting And Rejecting the Doctrine of Prior Entry. *Human Movement Science*, 9, 481-530.
- Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, 391(6666), 481-484.
- Hahn, S., & Kramer, A. (1998). Further Evidence for the Division of Attention Among Non-contiguous Locations. *Visual Cognition*, 5(1/2), 217-256.
- Heinze, H. J., Luck, S. J., Munte, T. F., Gos, A., Mangun, G. R., & Hillyard, S. A. (1994). Attention to adjacent and separate positions in space: an electrophysiological analysis. *Perception & Psychophysics*, 56(1), 42-52.

- Hillyard, S. A., & Anillo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings National Academy Society USA*, 95(3), 781-787.
- Hopf, J. M., Boehler, C. N., Luck, S. J., Tsotsos, J. K., Heinze, H. J., & Schoenfeld, M. A. (2006). Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proceedings National Academy Society USA*, 103(4), 1053-1058.
- Hopf, J. M., Luck, S. J., Girelli, M., Hagner, T., Mangun, G. R., Scheich, H., & Heinze, H.J. (2000). Neural sources of focused attention in visual search. *Cerebral Cortex*, 10(12), 1233-1241.
- Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, 394(6693), 575-577.
- Hughes, H. C., & Zimba, L. D. (1985). Spatial maps of directed visual attention. *Journal of Experimental Psychology Human Perception and Performance*, 11(4), 409-430.
- Hughes, H. C., & Zimba, L. D. (1987). Natural boundaries for the spatial spread of directed visual attention. *Neuropsychologia*, 25(1A), 5-18.
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, 43(3), 171-216.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40(10-12), 1489-1506.
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews: Neuroscience*, 2(3), 194-203.
- Jaskowski, P., van der Lubbe, R. H., Schlottbeck, E., & Verleger, R. (2002). Traces left on visual selective attention by stimuli that are not consciously identified. *Psychological Science*, 13(1), 48-54.
- Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, 43(4), 346-354.
- Kane, M. J., Poole, B. J., Tuholski, S. W., & Engle, R. W. (2006). Working memory capacity and the top-down control of visual search: exploring the boundaries of "executive attention". *Journal of Experimental Psychology Human Perception and Performance*, 32(4), 749-777.
- Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science*, 282(5386), 108-111.
- Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, 22(4), 751-761.
- Khayat, P. S., Spekreijse, H., & Roelfsema, P. R. (2006). Attention lights up new object representations before the old ones fade away. *Journal of Neuroscience*, 26(1), 138-142.
- Klotz, W., & Neumann, O. (1999). Motor activation without conscious discrimination in metacontrast masking. *Journal Experimental Psychology Human Perception and Performance*, 25, 976-992.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4(4), 219-227.
- Kramer, A., & Hahn, S. (1995). Splitting the Beam: Distribution of Attention Over Noncontiguous Regions of the Visual Field. *Psychological Science*, 6(6), 381-386.
- Kristjansson, A. (2000). In search of remembrance: evidence for memory in visual search. *Psychological Science*, 11(4), 328-332.
- Krose, B. J., & Julesz, B. (1989). The control and speed of shifts of attention. *Vision Research*, 29(11), 1607-1619.
- Kusunoki, M., Gottlieb, J., & Goldberg, M. E. (2000). The lateral intraparietal area as a salience map: the representation of abrupt onset, stimulus motion, and task relevance. *Vision Research*, 40(10-12), 1459-1468.
- LaBerge, D., & Brown, V. (1986). Variations in size of the visual field in which targets are presented: an attentional range effect. *Perception & Psychophysics*, 40(3), 188-200.
- LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, 96, 101-124.

- LaBerge, D., Carlson, R. L., Williams, J. K., & Bunney, B. G. (1997). Shifting attention in visual space: tests of moving-spotlight models versus an activity-distribution model. *Journal of Experimental Psychology Human Perception and Performance*, 23(5), 1380-1392.
- Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 23(11), 571-579.
- Luck, S. J., & Hillyard, S. A. (1994a). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31(3), 291-308.
- Luck, S. J., & Hillyard, S. A. (1994b). Spatial filtering during visual search: evidence from human electrophysiology. *Journal of Experimental Psychology Human Perception and Performance*, 20(5), 1000-1014.
- Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, 32(1), 4-18.
- McCormick, P. A. (1997). Orienting attention without awareness. *Journal of Experimental Psychology Human Perception and Performance*, 23(1), 168-180.
- McCormick, P. A., Klein, R. M., & Johnston, S. (1998). Splitting versus sharing focal attention: comment on Castiello and Umiltà (1992). *Journal Experimental Psychology Human Perception and Performance*, 24(1), 350-357.
- McMains, S. A., & Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron*, 42(4), 677-686.
- McMains, S. A., & Somers, D. C. (2005). Processing efficiency of divided spatial attention mechanisms in human visual cortex. *Journal of Neuroscience*, 25(41), 9444-9448.
- Moore, C. M., Egeth, H., Berglan, L. R., & Luck, S. J. (1996). Are attentional dwell times inconsistent with serial visual search? *Psychonomic Bulletin & Review*, 3, 360-365.
- Mordkoff, J. T., Yantis, S., & Egeth, H. E. (1990). Detecting conjunctions of color and form in parallel. *Perception & Psychophysics*, 48(2), 157-168.
- Morgan, S. T., Hansen, J. C., & Hillyard, S. A. (1996). Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proceedings of the National Academic Society U S A*, 93(10), 4770-4774.
- Motter, B. C. (1994). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neuroscience*, 14(4), 2178-2189.
- Muller, H. J., & Findlay, J. M. (1987). Sensitivity and criterion effects in the spatial cuing of visual attention. *Perception & Psychophysics*, 42(4), 383-399.
- Müller, M. M., & Hubner, R. (2002). Can the Spotlight of Attention be Shaped like a Doughnut? Evidence From Steady-State Visual Evoked Potentials. *Psychological Science*, 13(2), 119-124.
- Müller, M. M., Malinowski, P., Gruber, T., & Hillyard, S. A. (2003). Sustained division of the attentional spotlight. *Nature*, 424(6946), 309-312.
- Müller, M. M., Picton, T. W., Valdes-Sosa, P., Riera, J., Teder-Salecjarvi, W. A., & Hillyard, S. A. (1998). Effects of spatial selective attention on the steady-state visual evoked potential in the 20-28 Hz range. *Cognitive Brain Research*, 6(4), 249-261.
- Müller, M. M., Teder-Salecjarvi, W., & Hillyard, S. A. (1998). The time course of cortical facilitation during cued shifts of spatial attention. *Nature Neuroscience*, 1(7), 631-634.
- Müller, N. G., Bartelt, O. A., Donner, T. H., Villringer, A., & Brandt, S. A. (2003). A physiological correlate of the "Zoom Lens" of visual attention. *Journal of Neuroscience*, 23(9), 3561-3565.
- Müller, N. G., & Kleinschmidt, A. (2004). The attentional 'spotlight's' penumbra: center-surround modulation in striate cortex. *Neuroreport*, 15(6), 977-980.
- Müller, N. G., Mollenhauer, M., Rösler, A., & Kleinschmidt, A. (2005). The Attentional Field had a Mexican hat Distribution. *Vision Research*, 45, 1129-1137.
- Pan, K., & Eriksen, C. W. (1993). Attentional distribution in the visual field during same-different judgments as assessed by response competition. *Perception & Psychophysics*, 53(2), 134-144.

- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention & performance X* (pp. 531-556). Hillsdale, NJ: Erlbaum.
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology*, 109, 160-174.
- Pratt, J., & Quilty, L. (2002). Examining the activity-distribution model of visual attention with exogenous cues and targets. *Quarterly Journal of Experimental Psychology A*, 55(2), 627-641.
- Ress, D., & Heeger, D. J. (2003). Neuronal correlates of perception in early visual cortex. *Nature Neuroscience*, 6(4), 414-420.
- Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, 19(5), 1736-1753.
- Rizzolatti, G., Riggio, L., Dascola, I., & Umiltà, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, 25(1A), 31-40.
- Scharlau, I. (2004). Evidence for split foci of attention in a priming paradigm. *Perception & Psychophysics*, 66(6), 988-1002.
- Scharlau, I., & Ansorge, U. (2003). Direct parameter specification of an attention shift: evidence from perceptual latency priming. *Vision Research*, 43(12), 1351-1363.
- Scharlau, I., & Neumann, O. (2003a). Perceptual latency priming by masked and unmasked stimuli: evidence for an attentional interpretation. *Psychological Research*, 67(3), 184-196.
- Scharlau, I., & Neumann, O. (2003b). Temporal Parameters and the time course of perceptual latency priming. *Acta Psychologica*, 113, 185-203.
- Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive Psychology*, 47(4), 333-366.
- Shaw, M. L. (1978). A capacity allocation model for reaction time. *Journal Experimental Psychology Human Perception and Performance*, 4, 586-598.
- Shaw, M. L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal Experimental Psychology Human Perception and Performance*, 3, 201-211.
- Shiffrin, R. M., & Gardner, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology*, 93(1), 72-82.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. General Theory. *Psychological Review*, 82, 127-190.
- Shmuel, A., Yacoub, E., Pfeuffer, J., Van de Moortele, P. F., Adriany, G., Hu, X., & Ugurbil, K. (2002). Sustained negative BOLD, blood flow and oxygen consumption response and its coupling to the positive response in the human brain. *Neuron*, 36(6), 1195-1210.
- Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychological Science*, 12(3), 205-212.
- Shulman, G. L., Remington, R. W., & McLean, J. P. (1979). Moving attention through visual space. *Journal of Experimental Psychology Human Perception and Performance*, 5(3), 522-526.
- Slotnick, S. D., Hopfinger, J. B., Klein, S. A., & Sutter, E. E. (2002). Darkness beyond the light: attentional inhibition surrounding the classic spotlight. *Neuroreport*, 13(6), 773-778.
- Slotnick, S. D., Schwarzbach, J., & Yantis, S. (2003). Attentional inhibition of visual processing in human striate and extrastriate cortex. *Neuroimage*, 19(4), 1602-1611.
- Smith, A. T., Singh, K. D., & Greenlee, M. W. (2000). Attentional suppression of activity in the human visual cortex. *Neuroreport*, 11(2), 271-277.
- Smith, A. T., Williams, A. L., & Singh, K. D. (2004). Negative BOLD in the visual cortex: evidence against blood stealing. *Human Brain Mapping*, 21(4), 213-220.
- Somers, D. C., Dale, A. M., Seiffert, A. E., & Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academic Society U S A*, 96(4), 1663-1668.
- Standage, D. I., Trappenberg, T. P., & Klein, R. M. (2005a). A Continuous attractor neural network model of divided visual attention. Paper presented at the Proceedings of the International Joint Conference on Neural Networks, Montreal, Canada.

- Standage, D. I., Trappenberg, T. P., & Klein, R. M. (2005b). Modelling divided visual attention with a winner-take-all network. *Neural Networks*, 18(5-6), 620-627.
- Theeuwes, J. (1995). Abrupt luminance change pops out; abrupt color change does not. *Perception & Psychophysics*, 57(5), 637-644.
- Theeuwes, J., Godijn, R., & Pratt, J. (2004). A new estimation of the duration of attentional dwell time. *Psychonomic Bulletin & Review*, 11(1), 60-64.
- Titchener, E. M. (1908). *Lectures on the elementary psychology of feeling and attention*. New York: Macmillan.
- Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., & Dale A.M. (1998). The retinotopy of visual spatial attention. *Neuron*, 21(6), 1409-1422.
- Treisman, A. (1982). Perceptual grouping and attention in visual search for features and objects. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 194-212.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological Review*, 95(1), 15-48.
- Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology Human Perception and Performance*, 16(3), 459-478.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97-136.
- Tsal, Y. (1983). Movements of attention across the visual field. *Journal of Experimental Psychology Human Perception and Performance*, 9(4), 523-530.
- Ward, R., Duncan, J., & Shapiro, K. (1996). The Slow Time-Course of Visual Attention. *Cognitive Psychology*, 30(1), 79-109.
- Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1, 202-238.
- Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature*, 400(6747), 867-869.
- Woodman, G. F., & Luck, S. J. (2003). Serial deployment of attention during visual search. *Journal Experimental Psychology: Human Perception and Performance*, 29(1), 121-138.

Chapter 7

Reflections of complex spatial attention distributions in human visual cortex: preliminary results of an fMRI study exploring the “attentional landscape”

Based on: Peters, J. C., De Weerd, P., & Goebel, R. Reflections of complex spatial attention distributions in human visual cortex: preliminary results of an fMRI study exploring the “attentional landscape”. (*in preparation*)

Abstract

Although recent studies have suggested that spatial attention can be deployed in more complex configurations than a single unitary focus, the attainable degree of complexity remains unclear. The present functional magnetic resonance imaging (fMRI) study investigated attentional modulations in the visual cortex when attention was simultaneously allocated to either two small rectangles, two large rectangles, or one small and one large rectangle of three differently-sized, centrally aligned (outlined) rectangle pairs. Although preliminary, the obtained fMRI results suggest that attention could be exclusively deployed to the rectangles of interest. For example, the large and small rectangles could be attended in parallel without deploying attention to intervening regions, indicating that even when highly complex configurations are required the attentional system is highly efficient in allocating its limited resources. These results are in line with the emerging view that the distribution of attention is more appropriately characterized by a pliable landscape in which the weights of multiple locations can be modulated in parallel than by a serially moving unitary focus.

Introduction

In today's society, one is often confronted with complex visual scenes in which several objects need to be simultaneously attended, while others have to be ignored. While driving, other cars and traffic lights need to be attended, whereas neon advertising signs should to be ignored. Likewise, watching your children on a crowded playground or operating air traffic control equipment, requires an optimal allocation of one's limited attentional resources. Recent studies have shown that attentional configurations (i.e., the constellation of attended and non-attended areas) can be more complex than a single focused peak of attention as suggested by the attentional 'spotlight' (Posner, 1980) or 'zoom-lens' (Eriksen & St. James, 1986) model. Several psychophysical (Awh & Pashler, 2000; Bichot, Cave, & Pashler, 1999; Castiello & Umiltà, 1992; Hahn & Kramer, 1998; Kramer & Hahn, 1995; Scharlau, 2004) and neuroimaging (McMains & Somers, 2004; Müller, Malinowski, Gruber, & Hillyard, 2003) studies have suggested that multiple discrete parts of a visual scene can be attended in parallel (for review, see Chapter 6). In addition, other studies (Egley & Homa, 1984; Eimer 1999, 2000; Juola, Bouwhuis, Cooper, & Warner, 1991; Müller & Hübner, 2002; Somers, Dale, Seiffert, & Tootell, 1999) have shown that attention can be deployed in an annular fashion, while objects presented in the centre of this annulus are ignored. It would be highly efficient if attention could indeed be allocated to two non-contiguous regions or could be applied in an annular shape, given that due to limited attentional resources (e.g., Shaw & Shaw, 1977), a tradeoff exists between the size of the attended region and processing efficiency (e.g., Eriksen & St. James, 1986). In this way, valuable attentional resources would not be wasted by attending irrelevant or even distracting information at intervening areas (that would be encompassed by a spotlight or zoom-lens), thereby increasing the amount of attention that can be deployed to relevant areas. Indeed, a recent study by McMains and Somers (2005) indeed showed both improvements in performance as well as increased attentional enhancement in visual cortex when the region in between two task-relevant regions could be ignored. In addition, performance and attentional enhancement were similar when attending two adjacent or two segregated locations, suggesting that dividing attention is not associated with additional costs.

Although earlier work supported the notion that attention can be allocated in more complex configurations than a circular spotlight, the degree of complexity which can be achieved is not yet clear. For example, previous studies investigating divided attention used a paradigm in which attention could be split in two foci with a similar size. However, can attention also be allocated to two segregated regions of different sizes? The present fMRI experiment was performed to investigate the flexibility of the attentional system to establish complex attentional configurations. Participants had to detect the simultaneous appearance of two outlined rectangles (subsequently referred to as rectangles) with a pre-specified

size and color-orientation combination in a rapid serial visual presentation (RSVP) of small, medium and large pairs of uniquely colored rectangles (Figure 1). Note that the rectangles (of which the orientations varied over frames) almost resembled squares, which made the orientation discrimination very difficult. In addition, rectangles were only presented for 100 ms, which is too short to shift attention between rectangles in a serial fashion (which takes at least 150-200 ms; e.g., Duncan, Ward, & Shapiro, 1994; Müller, Teder-Salejarvi, & Hillyard, 1998; Ward, Duncan, & Shapiro, 1996; Weichselgartner & Sperling, 1987; Woodman & Luck, 2003; Chapter 6). Three attentional configurations were investigated. In the 'single attended region' conditions, participants covertly monitored either the large (*SINGLE-L*), or the small (*SINGLE-S*) rectangle pairs for target appearance. In the 'multiple attended regions' condition, participants had to detect a target defined by a combination of a small and a large rectangle (*MULTI*), necessitating the simultaneous monitoring of both relevant regions. These different attentional configurations were compared with a control condition, in which participants passively viewed the rectangles while focusing attention on the middle of the screen to monitor changes in the luminance of the fixation cross. The main focus of the present study was on the attentional distribution in the *SINGLE-L* and, especially, the *MULTI* condition. Since encompassing the large rectangles in an unitary attentional focus would require that attentional resources are distributed over a very large area (of at least 20° by 20°) in which distracting information would be present, it would be much more efficient if attention could be deployed in an annular fashion covering exclusively the large (*SINGLE-L*), or the large and small (*MULTI*) rectangles. Note that the latter condition required the parallel allocation of attention to regions segregated by large intervening areas. To investigate whether attention could be deployed in such complex configurations, patterns of activity enhanced by attention in the three conditions were examined in the visual cortex.

The use of rectangular stimuli offered several important advantages. Most importantly, since the ratio between the short and long side (further referred to as *side ratio*; short side = side ratio x long side) of each rectangle is very high, the whole rectangle needed to be attended to detect its orientation. This implies that the task cannot be solved locally by attending only parts of the rectangles. In addition, the jittering of the rectangle positions and the randomization of superimposing the colored rectangles prevented participants from using intersections between rectangles as an orientation cue. Another advantage of the present stimuli is that orientation discrimination difficulty can be adjusted by changing the ratio between the short and long sides of the rectangle (i.e., making the stimulus more rectangular or more square-like). The task difficulty could be optimized for each condition by solely introducing gradual changes in the side ratio, in the absence of large undesirable changes in stimulus appearance (see psychophysical study 1).

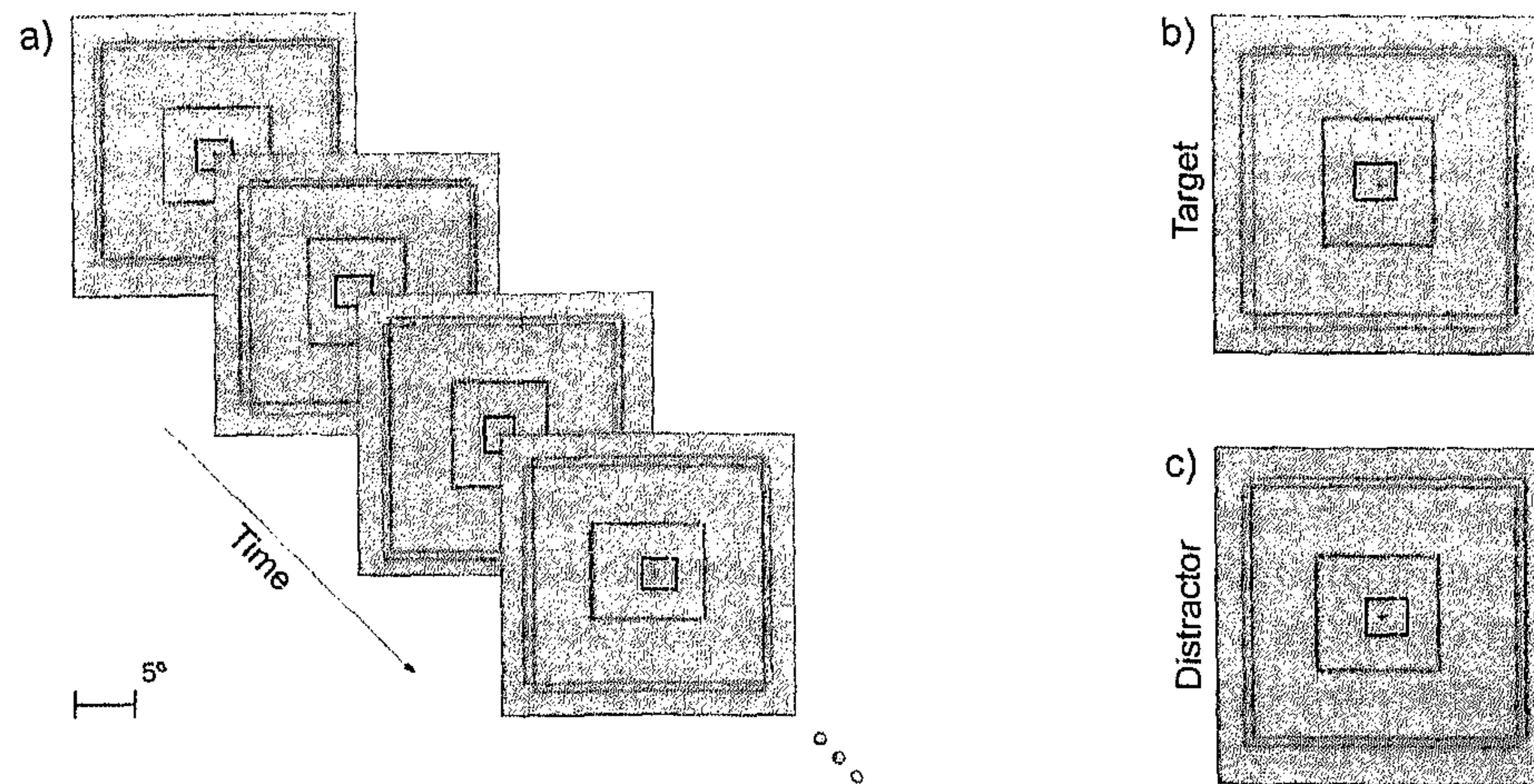


Figure 1. Schematic illustration of the task. a) Part of RSVP (stimuli are drawn to scale), in which participants either covertly monitored the two small rectangles (SINGLE-S), the two large rectangles (SINGLE-L), or one small and one large (MULTI) rectangle for targets. Target definitions for these three conditions could, for instance, be horizontal-green/vertical-red, horizontal-yellow/vertical-blue and horizontal-green/vertical-blue, respectively. Examples of b) target (i.e., horizontally oriented green and vertically oriented red rectangle) and c) distractor configurations in the SINGLE-L condition are shown on the right. Rectangle positions were jittered to prevent participants from using intersections of two rectangles as orientation indication. See text for details.

Two psychophysical experiments investigated the ability of participants to deploy attention in such complex configurations. In the first experiment, potential differences in task difficulty between the attentional configurations were estimated to equalize task load between the different conditions in the fMRI task as much as possible. A second experiment aimed to investigate the influence of the attentional configurations on probe detection in areas in between the small and large rectangles. The results of these experiments will be discussed first.

Psychophysical experiment 1: equalizing task difficulties between conditions by adjusting the side ratio

Orientation discrimination difficulty can be adjusted by changing the side ratio between the short and long sides of the rectangle. To explore potential differences in task difficulty between the attentional configurations used in the fMRI experiment, a staircase procedure was used in which the side ratios of the small and large rectangle pairs were independently adjusted according to performance.

Methods (Psychophysical experiment 1)

Design

Twelve right-handed healthy volunteers (4 males; mean age 20.7 years; standard deviation: 2.9 years) with normal or corrected-to-normal visual acuity were paid

for participation in this study. A stimulus display consisted of a small yellow, a small blue, a large red, and a large green outlined rectangle having a horizontal or vertical orientation presented on a grey background. The centre of the small ($3.0^\circ \times 2.4^\circ$) and large ($20.0^\circ \times 16.0^\circ$) rectangles varied randomly over trials, but was always within a radius of 0.22° (small) or 1.5° (large) around the fixation cross in the middle of the screen. To avoid overlap between rectangles of the same size, a constraint was implemented in the design such that the centers of rectangles with the same size and orientation should be at least 0.13° apart. Subjects covertly attended the two small rectangles (SINGLE-S), the two large rectangles (SINGLE-L), or one small and one large (MULTI) rectangle in separate blocks. For half of the participants (group 1), target specification of rectangle orientation and color was horizontal-yellow/vertical-blue for the SINGLE-S condition, horizontal-green/vertical-red for the SINGLE-L and horizontal-green/vertical-blue for the MULTI condition. For the other half of the participants (group 2), targets were specified as horizontal-blue/vertical-yellow, horizontal-red/vertical-green, and horizontal-blue/vertical-green, respectively. At the start of each block, subjects received an instruction concerning which specific stimulus configuration would be the target. To clarify, an instruction in a MULTI condition block for a participant in group 1 would be: “press the ‘target present’ button if the green large rectangle is horizontally oriented *and* the blue small rectangle is vertically oriented, else press the ‘target absent’ button.” Each trial within a block started with a ‘blank’ (i.e., only a fixation cross) screen (200 ms), followed by the presentation of the stimulus display. After onset of the stimulus, subjects were required to respond whether the target was present or absent, by pressing the ‘5’ or ‘6’ on the keypad of a keyboard with their right index or middle finger, respectively. A target was presented in 25% of the trials. Following the subsequent response interval (blank screen; 2.5 s), participants received auditory feedback (100 ms) about the correctness of their response (correct: high tone (1000 Hz); incorrect: low tone (500 Hz); miss: ringing tone).

Each block was composed of three parts: 1. A training part in which stimulus displays were presented for 200 ms (48 trials). 2. Another 48 trials with stimulus duration decreased to 100 ms. 3. In the third part, the side ratio initially was 0.80 (similar to the other parts). However, the side ratio changed over trials according to performance, in order to establish the side ratio threshold at which the participant could perform the task with an accuracy of 84%. To determine the threshold, a staircase procedure (Wetherill & Levitt, 1965) was used in which the side ratio of the rectangles (presented for 100 ms) was increased or decreased with 0.05 after each ‘reversal point’ (i.e., respectively 4 successive hits or 1 error). The staircase was terminated after 14 reversal points, after 96 trials (when the 14 reversal points were not reached), or when a subject reached the highest obtainable performance (i.e., when the next increase in difficulty would result in a side ratio of 1, which is a square with an undefined orientation). The average of the last six (of in total 14) reversal points was taken as the performance threshold. If subjects

had less than 9 reversal points, the threshold was regarded as non-stable and therefore not included in the group average. Each part of the block started with an instruction and a warm-up trial. The order of blocks and response buttons were counterbalanced over participants within the two groups. Within the first two parts of a block, each combination of orientations and colors of the four rectangles was presented three times (4 of the 16 combinations included a target configuration) in random order. Throughout the experiment, participants were reminded to fixate to the fixation cross during stimulus presentation and to respond as fast and accurately as possible. They could rest between blocks. Each participant was tested in a dimly-lit room while seated in front of a 17" CRT monitor with his or her head resting on a chin rest. The centre of the screen was at eye level and distance between the eyes and the monitor was 57 cm. Stimuli were presented and responses were recorded using the Presentation software package (version 9.20; Neurobehavioral Systems, San Francisco, USA).

Data analysis

Only correct responses that occurred within a reaction time (RT) window defined as the interval from two standard deviations below to two standard deviations above the subject-specific mean reaction time per condition were included in subsequent RT analyses. Mean RT and error percentage obtained of the second part of each block (in which the side ratio was constant) were submitted to a repeated-measures analysis of variance (ANOVA) with two within-subject factors (block [SINGLE-L, SINGLE-S, MULTI] and target presence [present, absent]) and one between-subject factor (target specification [group 1, group 2]). Greenhouse-Geisser correction was applied for all p-values.

Results & Discussion (Psychophysical experiment 1)

There was a main effect of condition on RT ($F(2,20) = 15.5$, $p = .001$) and on error percentages ($F(2,20) = 8.5$, $p < .05$). Post-hoc comparisons showed that RT in the MULTI condition (mean(\pm standard error); 928(\pm 93) ms) was slower than in the SINGLE-S (687(\pm 60) ms; $t(11) = 5.0$, $p < .01$) and SINGLE-L (683(\pm 58) ms; $t(11) = 3.8$, $p = .000$) condition. In addition, more errors were made in the MULTI condition (14.6(\pm 3.8)%) compared to the SINGLE-S (4.4(\pm 1.4)%) condition ($t(11) = 3.7$, $p < .003$), but not compared to the SINGLE-L (5.5(\pm 1.5)%) condition. The results of the third (staircase) part also indicated that the MULTI condition was more difficult than the other two. Subjects performed very well in the SINGLE-L and SINGLE-S conditions: In both conditions, participants needed on average 5 reversal points to reach the highest obtainable performance (i.e., a side ratio of 95%) and in both conditions only two subjects did not reach this ceiling performance. Since none of the participants had more than 8 reversal points in both conditions, we averaged the last reversal point of all subjects, resulting in a side-ratio threshold of 0.93 for the SINGLE-L and 0.94 for the SINGLE-S condition. In contrast to the SINGLE conditions, only 4 of the 12 subjects reached

ceiling performance in the MULTI condition. The average side ratio of the last 6 reversal points (of participants with more than 8 reversal points) was 0.84. Since a pilot study ($n=2$) suggested that even extensive training of the MULTI condition cannot compensate for the increased difficulty compared to the other conditions, the obtained side-ratio thresholds were used in subsequent studies to equalize task load between conditions. However, since the fMRI task was more demanding than the present task, due to the rapid serial presentation and the inclusion of a medium pair of rectangles, slightly lower side ratios were used: 0.90 and 0.80 for the SINGLE and MULTI conditions, respectively.

Psychophysical experiment 2: estimating attentional allocation to intermediate regions using a probe detection paradigm

A probe detection task was combined with the task of the previous experiment to estimate potential differences in the amount of attention deployed to intermediated regions. Several neuroimaging studies have suggested that attention is highly spatially selective (e.g., Brefczynski, & DeYoe, 1999; McMains & Somers, 2004) even if the area has an annular shape (Müller & Hübner, 2002). Therefore, attention is expected to be confined to the small or to the large rectangles in the SINGLE-S and SINGLE-L condition, respectively.

On some of the trials, a small, non-salient probe appeared somewhere in the region in between the small and large rectangles. If participants are not able to divide attention over two non-contiguous locations in the MULTI condition but also deploy attention to the intervening area, the facilitated processing of stimuli in that area would lead to faster probe detection in the MULTI compared to the SINGLE conditions.

Methods (Psychophysical experiment 2)

Design

Twelve right-handed healthy volunteers (2 males; mean age 20.3 years; standard deviation: 1.6 years) with normal or corrected-to-normal visual acuity were paid for participation in this study. The stimuli, experimental design and procedure were identical to the first psychophysical study, unless stated otherwise. Results of this first study were used to equalize task load between the different conditions. The small and large rectangles had a size of $3.0^\circ \times 2.4^\circ$ and $20.0^\circ \times 16.0^\circ$, respectively, (side ratio = 0.80) for the MULTI condition. In the other two conditions, sizes were $3.0^\circ \times 2.7^\circ$ and $20.0^\circ \times 18.0^\circ$ (side ratio = 0.90) for the small and large rectangles. Since different target specifications did not give rise to differential performance in the staircase experiment, one target specification was used in the present experiment (i.e., horizontal-yellow/vertical-blue for the SINGLE-S condition, horizontal-green/vertical-red for the SINGLE-L and

horizontal-green/vertical-blue for the MULTI condition). To investigate whether the allocation of attention to the appropriate regions in visual space was facilitated when these regions were cued throughout the experiment, participants were randomly divided over two groups. Two grey concentric rings outlining the areas in which the rectangles could appear, served as background for one group, whereas a less informative fixation cross was presented throughout the experiment to the other group.

In the stimulus display, a non-salient probe (consisting of 4 red-green (50%) or yellow-blue (50%) squares) could appear on 8 locations (the corners and midpoints) of a virtual outlined square located precisely in between the large and small rectangle area. The small ($0.26^\circ \times 0.26^\circ$) probe appeared simultaneously with the rectangles on the screen in order not to produce a pop-out effect. If subjects detected this imperative probe in the display, they should not perform the target detection task, but instead press '5' on the keypad (right index finger). Otherwise, participants indicated target occurrence by pressing the '4' and '6' on the keypad with their right index or middle finger. Subjects received one block in which 6.25% and one block in which 25% of the trials a probe was presented for each condition, resulting in 6 blocks in total. Probe occurrence percentage was varied over blocks to influence the strategy of participants, assuming that the main task would receive more attention than probe detection in the 6.25% compared to the 25% probe probability variant. Trials in which the probe was present or absent varied randomly within blocks. Participants were informed about the target definition and the probe occurrence percentages at the start of each block. Each combination of orientations and colors of the four rectangles was presented four times (4 of 16 combinations included a target configuration) in random order within each block. The order of blocks and response buttons were counterbalanced over participants within the two groups. Participants trained each condition with slightly more rectangular stimuli (MULTI: $3.0^\circ \times 2.25^\circ$ and $20.0^\circ \times 12.0^\circ$; side ratio = 0.75; SINGLE: $3.0^\circ \times 2.55^\circ$ and $20.0^\circ \times 17.0^\circ$; side ratio = 0.85) that were presented for 200 ms.

Data analysis

Similar to the previous experiment, only correct responses that occurred within a RT window defined as the interval from two standard deviations below to two standard deviations above the subject-specific mean reaction time per condition were included in subsequent RT analyses. Two participants (one of each group) were excluded due to high mean RT or low accuracy. Mean RT and error percentage of trials in which a probe was presented were submitted to a repeated-measures ANOVA with three within-subject factors (condition [SINGLE-L, SINGLE-S, MULTI], probe probability [25%, 6.25%] and probe color [green-red, yellow-blue]) and one between-subject factor (fixation stimulus [2 concentric rings, fixation cross]). Performance scores for the other trials were analyzed with a repeated-measures ANOVA with similar factors as the ANOVA of the probe task,

except that the probe color was replaced by a target presence (present, absent) factor. All reported p-values were Greenhouse-Geisser corrected.

Results & Discussion (Psychophysical experiment 2)

Results suggested that the presentation of a cue indicating the areas that should be attended, instead of a less informative fixation cross, did not influence performance on the main task nor on the probe detection task (all p-values > .1). When a probe was presented, participants responded on average after 694(±22) ms, being accurate in 97.8(±0.4)% of the cases. Only a main effect of probe probability was obtained in the RT ($F(1,1) = 12.5$; $p = .008$) and error ($F(1,1) = 10.2$; $p = .013$) analyses, revealing that responses were slower but less erroneous in 6.25% compared to 25% probe probability blocks. Most importantly, RT ($p > .9$) and error percentages ($p > .5$) did not differ between conditions. In trials in which no probe was presented, mean RT and accuracy were 852(±28) ms and 89.9(±1.1)%, respectively. A main effect of target presence on RT was observed ($F(1,1) = 9.5$; $p = .015$) indicating that subjects were slower in on target present than target absent trials. In addition, a main effect of condition ($F(2,2) = 4.7$; $p = .033$) was found, suggesting that responses in the MULTI condition (937(±56) ms) were slower than in the other conditions (SINGLE-L: 820(±32) ms; SINGLE-S: 800(±37) ms). However, none of the post-hoc comparisons between the different conditions survived Bonferroni correction. A main effect for probe probability ($F(1,1) = 10.0$; $p = .013$) and target presence ($F(1,1) = 15.4$; $p = .004$) on accuracy was obtained. Since the interaction between these two factors also reached significance ($F(1,1) = 8.8$; $p = .018$), effects were estimated per level of the two factors. Post-hoc comparisons revealed that more misses were observed in blocks with 25% compared to 6.25% probe probability. No other significant effects were obtained.

The equal probe detection performance in the different conditions indicated that in the different conditions an equal amount of attention was deployed to the region in between the large and small rectangles. Under the assumption of a precise allocation of attention to the area of interest in the single attention conditions, these results suggest that in the MULTI condition attention was divided over the non-contiguous regions in which the large and small rectangles appeared, without deploying attention to the irrelevant intervening region.

fMRI experiment: reflections of a complex attentional configuration in early visual cortex

Although the behavioral results of experiment 2 suggested that no attention was deployed to the regions in between the small and large rectangle pair, no detailed information was offered concerning the attentional distribution in each condition.

To gain more insights into these different distributions, the high spatial resolution of fMRI was exploited to conduct a fine-grained analysis of the attentional enhancements in retinotopic visual areas.

Experimental design

Two healthy, right-handed participants (1 male; age: 26 and 28 years) with normal or corrected to normal visual acuity volunteered in the present study. These two participants will be further referred to as P1 and P2. Similar to the psychophysical experiments, the study was approved by the local ethical committee.

Main experiment

Since the present study aimed to focus on the capability of the attentional system to have sustained rather than transient complex configurations, the feasibility of a sustained attentional configuration was optimized by introducing several features in the experimental design. Firstly, stimuli were presented in rapid succession instead of one stimulus per trial. Secondly, the interstimulus interval (ISI; Gaussian distribution with an average of 800 ms, ranging from 300 to 1300 ms, in discrete steps of 200 ms) as well as the intertarget interval (average of 3.9 s; ranging from 1.8 – 6.3 s) were strongly jittered, resulting in highly unpredictable stimulus and target occurrences. Thirdly, to facilitate the maintenance of the attentional configuration during the ISI, three flickering (5 Hz; duty-cycle 50%) light grey squares with the average size of the different rectangle pairs were presented, centered at the middle of the screen. Similar to psychophysical experiment 2, the frames had a slightly more rectangular shape in the MULTI compared to SINGLE conditions to equalize task difficulty as much as possible. This physical difference in stimulus appearance was that subtle, however, that it elicited no differential activation patterns in the visual areas, as suggested by a pilot session. Furthermore, an additional pair of medium sized rectangles was added to the stimulus display, to equalize sensory stimulation in the areas of interest (i.e., the area in which large and the area in which small rectangles are presented and the area in between them) as much as possible.

Subjects covertly monitored the two small (SINGLE-S), or the two large (SINGLE-L) rectangles, or one small and one large (MULTI) rectangle for targets in a RSVP (for an example, see Figure 1). The RSVP lasted for 40 s and consisted of 44 displays (100 ms) of six uniquely colored, vertically or horizontally oriented rectangles. The small yellow and blue (MULTI: $3^\circ \times 2.3^\circ$; SINGLE: $3^\circ \times 2.6^\circ$), the medium turquoise and purple (MULTI: $9^\circ \times 7.2^\circ$; SINGLE: $9^\circ \times 8.1^\circ$), and the large red and green (MULTI: $20^\circ \times 16.5^\circ$; SINGLE: $20^\circ \times 18.5^\circ$) rectangle pairs were isoluminant (47 cd/m^2). The centre of the rectangles varied randomly over displays but was always located within a radius of 0.22° (small), 0.66° (medium), or 1.5° (large) around the fixation cross in the middle of the screen. To avoid overlap between rectangles of the same size, a constraint was implemented in the design such that the centers of rectangles with the same size and orientation should

be at least 0.17° apart. Each run consisted of a MULTI, a SINGLE-L and a SINGLE-S block, in which subjects covertly monitored an RSVPs followed by a fixation periods of 20 s three times in a row. At the start of each RSVP, participants received a cue indicating which rectangles should be attended and an additional reminder on the corresponding target specification. Targets were defined as horizontal-yellow/vertical-blue for the SINGLE-S condition, horizontal-green/vertical-red for the SINGLE-L and horizontal-green/vertical-blue for the MULTI condition. Participants were instructed to respond as fast and accurately as possible, by pressing the button of a buttonbox with their right index finger, whenever they detected a target (which was present in 18.18% of the stimulus displays, i.e., 24 times per block). To match stimulus appearance between the main and the control experiment (see next paragraph), the luminance of the fixation cross was increased for 100 ms 24 times per block. In the main experiment, this change did not need to be reported by the subjects. All stimuli were presented on a grey background and a fixation cross was presented throughout the experiment. The order of blocks was counterbalanced over runs and randomized across subjects. Each combination of orientations and colors of constituting a target configuration was presented equally often within a block, which was also the case for non target configurations. Target and distractor configurations were pseudo-randomized within blocks. Throughout the experiment, participants were reminded to fixate to the fixation cross during stimulus presentation and to respond as fast and accurately as possible.

Control experiment

Display conditions in the control experiment were identical to the experimental runs, except that the cue presented in advance of each block indicated to perform the luminance task and two non-informative grey squares were presented. Participants were instructed to press the button corresponding to their right index finger, whenever they detected a change in the luminance of the fixation cross (which occurred 24 times per block for 100 ms).

Behavioral training session

All participants took part in a behavioral training session prior to scanning. To equalize task difficulty as much as possible, participants were trained until they reached a performance of 73.5% in all three conditions. The experimental design was identical to the design of the experimental runs of the fMRI session, except that the fixation period in between RSVPs was reduced from 20 s to 8 s. In addition, due to a different stimulus presentation set-up (i.e., a CRT monitor instead of a LCD projector), the hue and luminance of some rectangles had to be slightly adjusted to maintain isoluminance between rectangle pairs. In contrast to the fMRI session, participants received information, whether a response was (high tone) or was not (ringing tone) detected in a response interval (300 -1600 ms) after

every target presentation. Responses outside this interval were identified as false alarms (low tone).

The training session consisted of four different parts: First, participants were familiarized with the task using a version of the task in which a target/non-target response was required for every stimulus. Then, participants performed a MULTI, SINGLE-L, and SINGLE-S block, each containing 3 RSVPs. The average ISI was 1000 ms (Gaussian, ranging from 500 to 1500 ms in discrete steps of 200 ms), resulting in an average intertarget interval of 4.7 s (ranging from 2.2 – 7.7 s). Subsequently, the ability to maintain the desired attentional configurations was trained, using a staircase procedure in which the ISI and, correspondingly, the intertarget interval was adapted according to performance. For each condition, RSVP/fixation period cycles were repeated within each block until 73.5% accuracy was obtained with an ISI and intertarget interval identical to those which would be used in the fMRI session. Finally, eye movements were monitored in the last run consisting of a MULTI, SINGLE-L, and SINGLE-S block (3 RSVPs) with an ISI and intertarget interval similar to the fMRI session. The setting in which the participant was tested was similar to that of the other psychophysical studies, except that eye movements were monitored using an Eyelink tracker with a sampling rate of 250 Hz (SensoMotoric Instruments GmbH, Teltow, Germany). Eye movements were categorized as saccades when the movement velocity exceeded 30°/s or when the movement acceleration was higher than 8000°/s².

Experimental design: localizer experiments

Retinotopic mapping

The early visual areas (V1, V2, V3, VP, V3a and V4v; Felleman & Van Essen, 1991) were defined for each subject based on retinotopic mapping scans (e.g., Sereno et al., 1995). Eccentricity and polar maps were acquired using, respectively, a gradually expanding annulus and a gradually rotating wedge composed of radial black and white checkerboards that reversed contrasts at 8 Hz. Each retinotopic mapping run consisted of eight repetitions (96 s each) of full expansions and full rotations, respectively.

‘Stimulus display localizer’

Visual regions, representing the parts of the visual field in which rectangle stimuli of the main task were presented, were localized using outlined squares composed of flickering (8 Hz) black-white checkerboards. The size and border thickness of the outlined square was defined by the maximum jitter of the rectangle positions. The outlined square had a size of 21.79° x 21.79° for the large, 10.01° x 10.01° for the medium, and 3.57° x 3.57° for the small rectangle area. The border thickness of each of the squares was 3.86°, 1.91°, and 0.87°, respectively. The run consisted of 4 blocks in which two repetitions of three 30 s blocks (a small, medium and large frame block in random order) were presented. Blocks and fixation periods

(20 s screen with fixation cross) were interleaved (and the run started and ended with a fixation period).

Image acquisition

Echo-planar images (T2-weighted; 128x128 imaging matrix, 28 slices, voxel size: 2 x 2 x 2 mm³, no gap, TE/TR= 30/2000 ms, FA= 90°) covering the visual cortex were collected on a 3-T Siemens Magnetom Allegra head-scanner (Siemens Medical Systems, Erlangen, Germany) using a standard head coil. Functional data were aligned to a T1-weighted high-resolution anatomical image (MDEFT sequence [Deichmann, Schwarzbauer, & Turner, 2004]; TE/TR = 2.4/7.92 ms; FA = 15°, voxel size = 1 mm³). Participants viewed the stimuli, projected onto a frosted screen using a liquid crystal display projector (PLC-XT11-16, Sanyo North America Corporation, San Diego, USA), via a mirror mounted to the head coil. Distance between the eyes and the screen was 47.5 cm. Stimuli were presented and responses were recorded using the Presentation software package (version 9.20; Neurobehavioral Systems, San Francisco, USA). Onsets of RSVPs and fixation periods were synchronized with MR acquisition. The main experiment consisted of two control and six experimental runs of 280 volumes each. To avoid interference from the experimental task, control runs were presented first. In addition, two stimulus display localizer runs (410 volumes each) and a polar angle and eccentricity mapping (268 volumes) was performed in the same or a different scanning session. Each participant took part in a separate training session, in which eye movements were monitored.

fMRI data analysis

Preprocessing of the individual datasets included slice scan time correction, linear trend removal, high-pass filtering, three-dimensional motion correction, transformation into Talairach space (Talairach & Tournoux, 1988), and cortex reconstruction, inflation and flattening as implemented in the BrainVoyager software package QX version 1.7.9 (Brain Innovation, Maastricht, the Netherlands). The first four volumes of each run were discarded to remove T1 saturation effects.

For each subject, Regions-of-Interest (ROIs) in the occipital cortex were defined as the retinotopic representations of the area in which the small, medium, or large rectangles could appear according to the results of the independent stimulus display localizer. This resulted in respectively a *small ROI*, *medium ROI*, and *large ROI* of the upper and lower visual field representation in each hemisphere. In addition, based on polar mapping results ROIs were defined for each visual area (V1, V2, V3, VP, V3a, and V4v). By combining information from the ROIs based on the stimulus display localizer and the polar mapping, additional ROIs were obtained in which the small, medium and large ROI were specified for each visual area. To assess attentional modulations within specified retinotopic regions, multiple regression analysis was performed for each ROI. Condition

effects were modeled using a boxcar function, convolved with a theoretical Two Gamma hemodynamic response function.

Results

Behavioral results

Responses occurring between 300 and 1600 ms after target onset were classified as correct responses, while responses outside this interval were designated as false alarms. Only correct responses were included in reaction time (RT) computations. Overall, participants responded after 696 (standard error: ± 25) ms with an accuracy of 79.0(± 3.2)%. Mean RT and accuracy of responses in the first (0-20 s; RT: 688(± 21) ms; accuracy: 79.8(± 3.0 %) and second half (20-40 s; RT: 704(± 30) ms; accuracy: 78.3(± 3.5 %) of the RSVP seemed not to differ. Furthermore, mean RT appeared to be higher whereas accuracy was lower for the SINGLE-L (RT: 729(± 53) ms; accuracy: 72.0(± 6.9 %) and MULTI (RT: 729(± 9) ms; accuracy: 72.7(± 2.7 %) compared to the SINGLE-S (RT: 631(± 13) ms; accuracy: 92.3(± 0.2 %) condition. However, due to the low number of participants, repeated-measures analyses of variance with trial phase (early, late) and condition (SINGLE-L, SINGLE-S, MULTI) as factors, revealed no significant effects.

fMRI results

The retinotopic representations of the area in which the small, medium, or large rectangles could appear, were identified for each subject using the independent 'stimulus display localizer', resulting in *small*, *medium*, and *large* ROIs. As illustrated in the outer sections of Figure 2, the large (red patches), medium (purple patches), and small (green patches) ROIs are arranged along a foveal-peripheral gradient. The overlap in neural activity between sensory stimulation of - and covert attention to - a region in the visual field was estimated by superimposing the ROIs on a statistical map contrasting activation in the SINGLE-L (orange colors) and SINGLE-S (blue colors) condition (inner sections of Figure 2). The inner sections of Figure 2 show that although sensory stimulation of the region representing the area in which the large and small rectangles appears (i.e., the larger and small ROI) strongly overlaps with attending respectively the large and small rectangles in the RSVP, attention to the large (orange) compared to the small (blue) rectangles also activated parts of the (especially left) occipital cortex representing the medium rectangles (i.e., the medium ROI; marked by the black-white dashed lines).

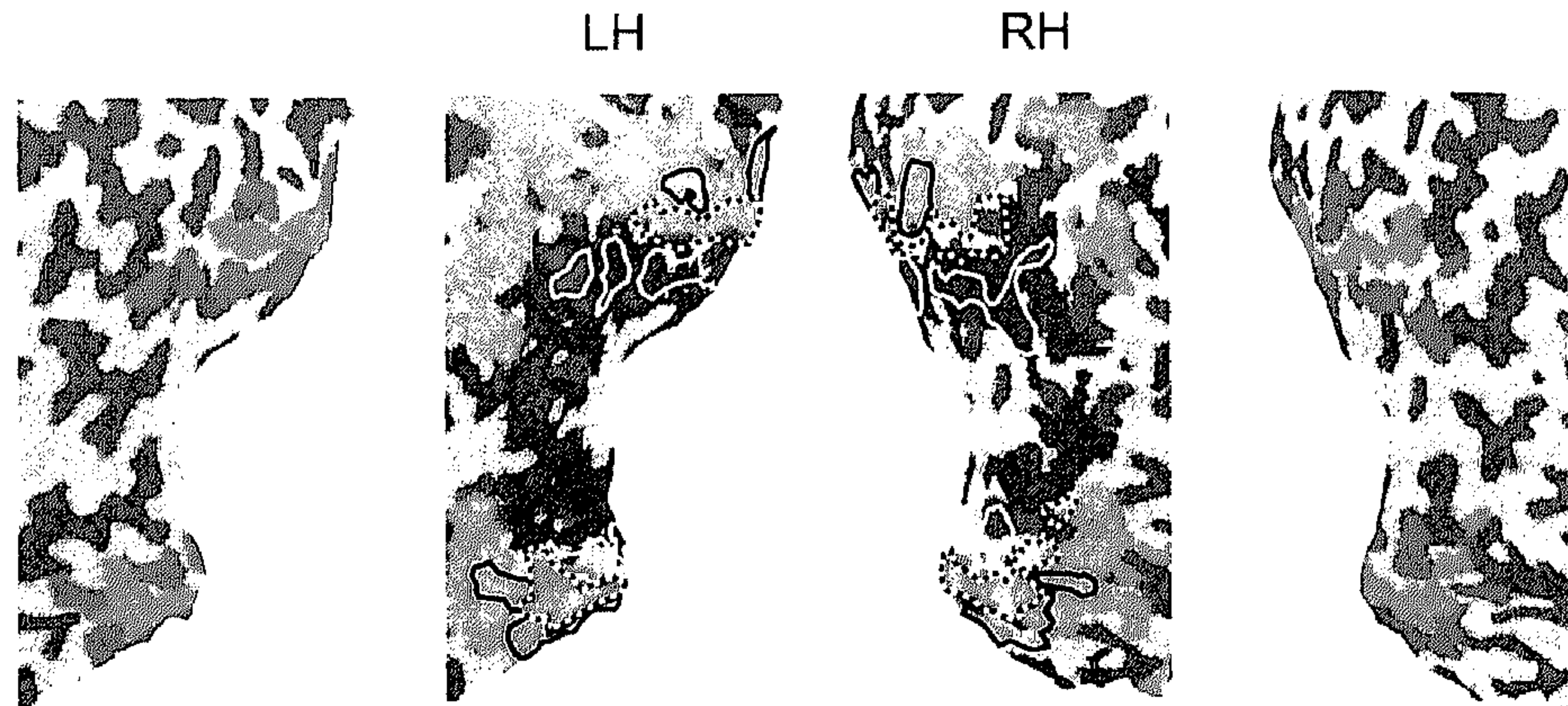


Figure 2. ROIs and statistical maps projected on the flattened representation of a participant's (P1) grey-white matter boundary surface (LH: left hemisphere; RH: right hemisphere; dark grey: concave light grey: convex cortical surface folding). Outer sections: ROIs representing the brain regions that respond strongest to the large (red), medium (purple), and small (green) outlined checkerboard squares of the stimulus display localizer. Inner sections: Statistical map contrasting activation in the SINGLE-L (orange) and SINGLE-S (blue) condition. Black black-white, and white lines mark the boundaries of respectively the large, medium, and small rectangle ROI (similar to the ROIs shown in the outer sections), demonstrating the coincidence of sensory stimulation of an covert attention to a region in the visual field.

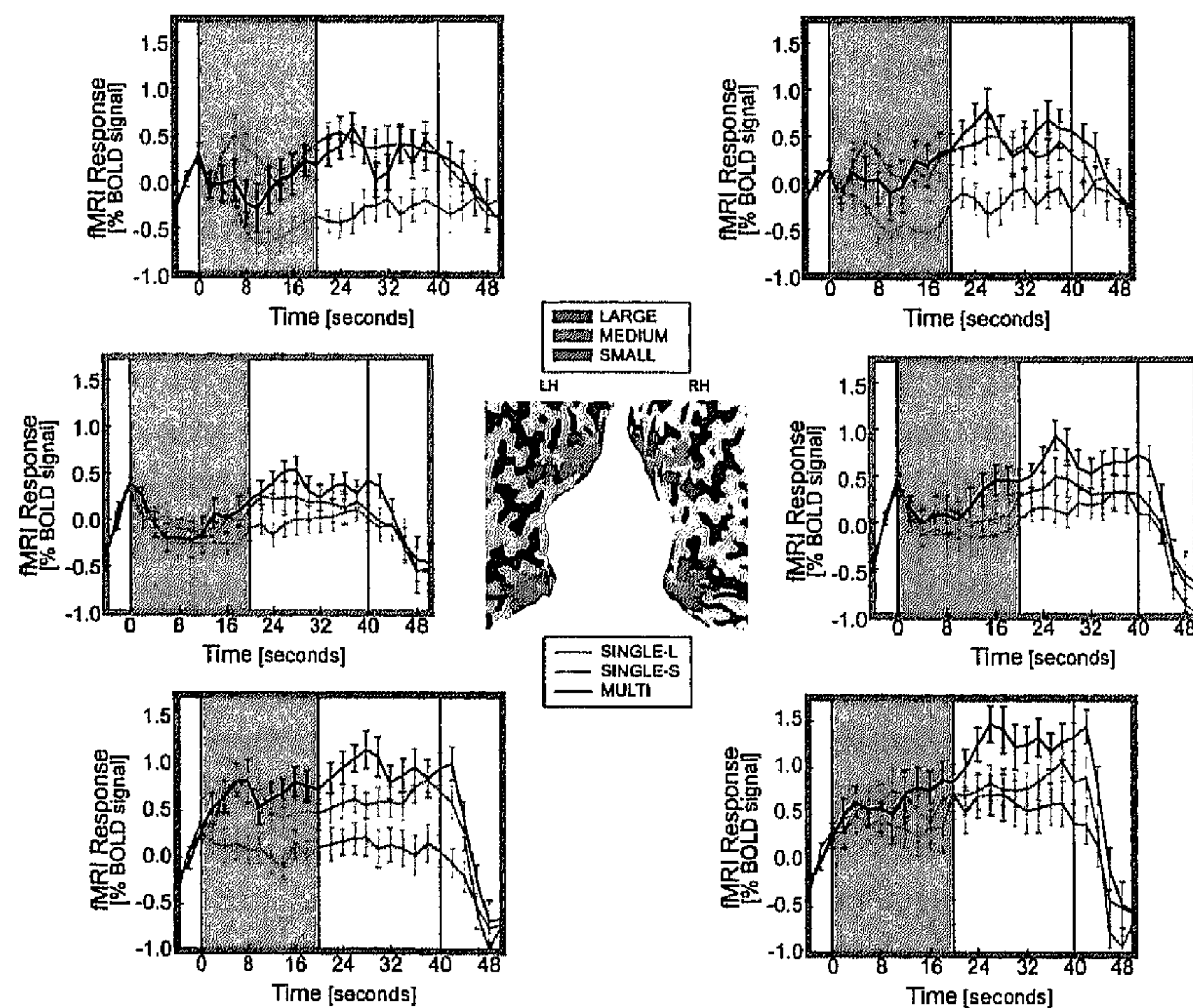


Figure 3. Event-related responses to the different conditions in the small, medium, and large ROI in the left (LH) and right (RH) hemisphere (shown for P1). Note that regions within the medium ROI that showed strong enhanced activity for the SINGLE-L compared to SINGLE-S condition were excluded. The dark and light grey boxes represent the first (early) and second (late) half of the RSVP. Error bars indicate standard errors of the mean (SEM).

Consistent with the statistical maps shown in Figure 2, average event-related responses of the small ROI (Figure 3; lower plots) appeared to be higher in the SINGLE-S and MULTI condition compared to the SINGLE-L condition. Likewise, activity in the large ROIs (Figure 3; upper plots) seemed to be enhanced in the SINGLE-L and MULTI condition compared to the SINGLE-S condition. The medium ROIs (Figure 3; middle plots), on the other hand, appeared not to differentially respond to the two SINGLE conditions. However, although responses in the MULTI and SINGLE conditions were initially similar, they appeared to gradually diverge over the course of the RSVP. Therefore, the first ('early phase') and second ('late phase') half of each trial was separately examined in the analysis that was subsequently performed for each ROI of each participant (Figure 4). Note that in this analysis attention conditions were contrasted against each other instead of against the control condition, since rectangles appeared not to be passively viewed in the control condition. This was suggested by visual inspection of the corresponding event-related averages indicating that enhancements were not confined to the fixation cross (e.g., activation in the small ROI was enhanced in the control compared to the SINGLE-S condition).

Aside from some incidental exceptions, the results were overall in line with our hypotheses (Figure 4). That is, activation in the small ROI was enhanced in the SINGLE-S compared to the SINGLE-L condition, whereas the reversed pattern was revealed in the large ROI. In addition, the MULTI condition elicited higher activation compared to the SINGLE-L (and SINGLE-S) condition in the small (respectively large) ROI. Finally, activity in the medium ROI did not differ between the SINGLE-S and SINGLE-L condition. Together, these results indicate that the small and large rectangles can be selectively attended. However, the primary focus of the present study was on potential diverging responses for the SINGLE and MULTI conditions in the medium ROI. The different results for the early and late phase suggested that activation patterns changed over time. In the early phase, activation strength did not differ between the three conditions. In contrast, the MULTI condition did elicit an increased response compared to the SINGLE conditions in the late phase. However, also the small and the large ROI showed an increased difference in the late compared to the early phase between the MULTI condition and the condition in which respectively the small (SINGLE-L) and large (SINGLE-S) rectangles were not attended. The increased activity in all three ROIs over the course of a MULTI condition RSVP suggested that diffuse changes in arousal, affecting activity in the entire visual cortex, might have contributed to the differential medium ROI responses in the late phase.

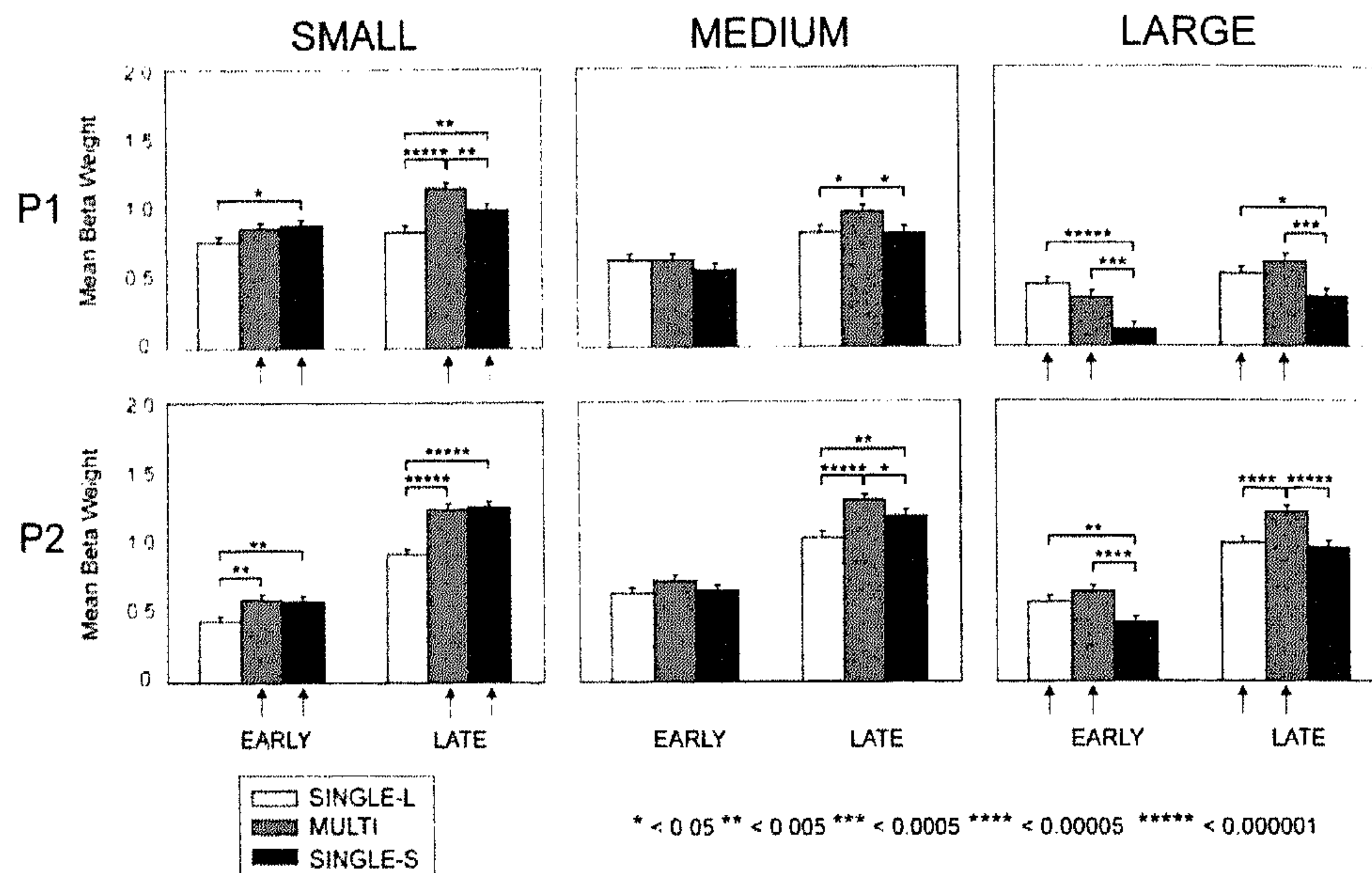


Figure 4. Mean beta weights of early (left clusters) and late (right clusters) responses in small (left column), medium (middle column), and large (right column) ROIs (averaged over hemispheres; see legend Figure 3 for details) shown for participant 1 (upper row) and 2 (lower row). Arrows indicate conditions in which the rectangles corresponding to the ROI should be attended. Error bars indicate SEM.

To examine activation patterns across visual areas, additional analyses were carried out in ROIs created by subdividing the small, medium, and large ROI according to the subject-specific retinotopic mapping results. Overall, findings suggested that the differences between the conditions were quite similar across visual areas (Figure 5).

Discussion

The present study aimed to extend recent findings showing that attention can be allocated in more complex configurations than suggested by the spotlight or zoom-lens model of attention. Although preliminary, the results of the fMRI experiment suggest that, in line with the results of psychophysical experiment 2, attentional arrangements can be highly complex.

Activity in the visual cortex was examined while subjects covertly attended to two small rectangles (SINGLE-S condition), two large rectangles (SINGLE-L), or one small and one large rectangle (MULTI) of three differently-sized, centrally aligned rectangle pairs. Consistent with the retinotopic organization of visual spatial attention (Brefczynski & DeYoe, 1999; Somers et al., 1999; Tootell et al., 1998), activity in the visual regions retinotopically

corresponding to the position of the small rectangles (small ROI) was enhanced when attention was deployed to one (in the MULTI condition) or to both (SINGLE-S) of the small rectangles. Likewise, the large ROI showed increased responses when one (MULTI) or both (SINGLE-L) of the large rectangles had to be attended.

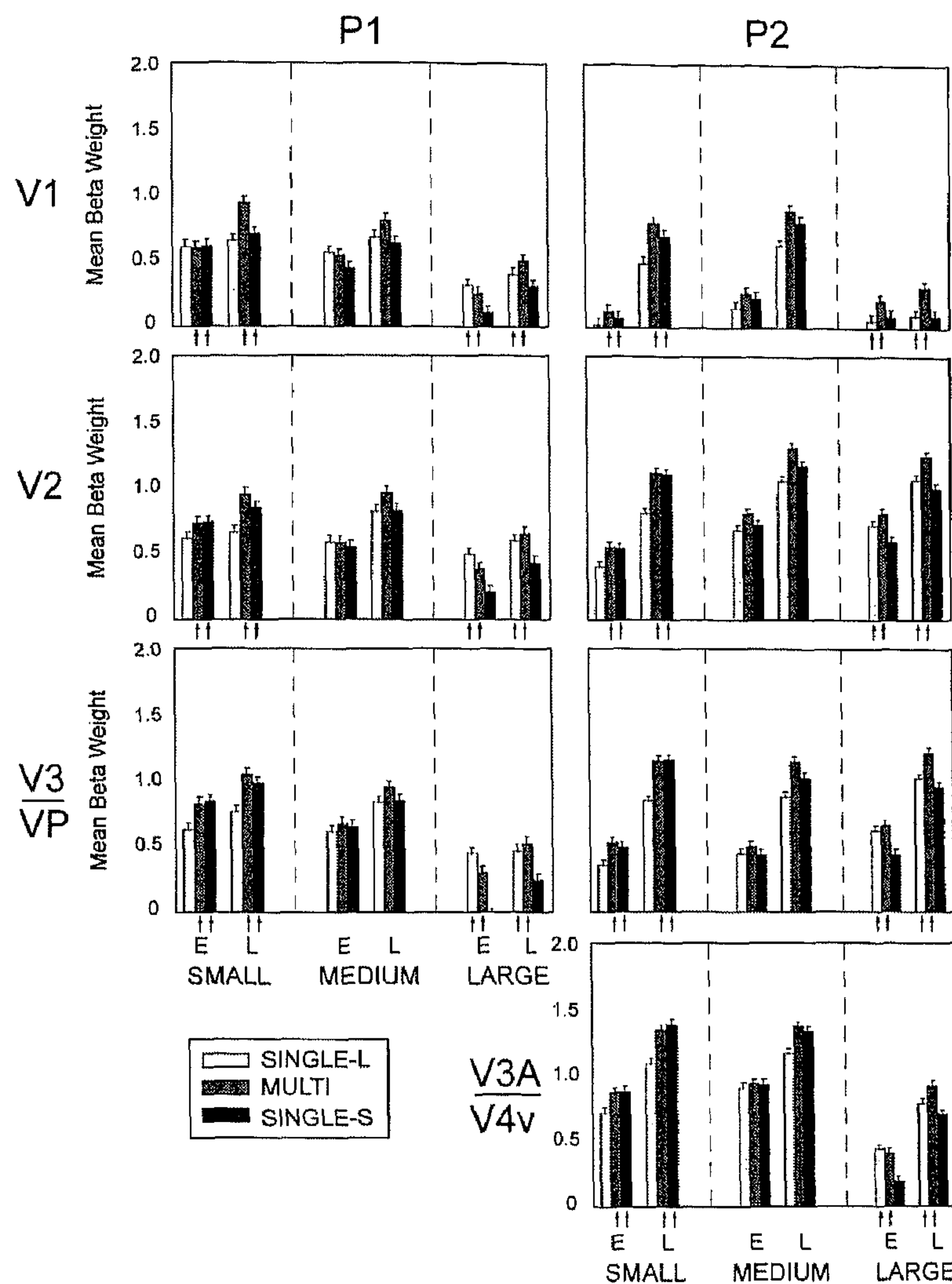


Figure 5. Mean beta weights of early (E) and late (L) responses in small, medium, and large subject-specific (columns) ROIs, separately shown for each mapped visual area (averaged over upper and lower visual field representations and over hemispheres). Arrows indicate conditions in which the rectangles corresponding to the ROI should be attended. Error bars indicate SEM.

In contrast to these specific attentional modulations in the small and large ROI, activity in the medium ROI did not differ between the SINGLE-S and SINGLE-L

condition. Since the control condition could not be used as 'passive viewing' baseline, these preliminary results reflect the relative rather than the absolute attentional distribution in the different conditions. Nevertheless, in line with previous fMRI studies revealing that attentional enhancements are highly spatially selective (e.g., Tootell et al., 1998) and can be shaped in an annular distribution (Somers et al., 1999), the present findings seem to suggest that attention can be confined to the large and small rectangle area in the SINGLE-L and SINGLE-S condition, respectively.

The primary research question of the present study, however, was whether attention could be divided between the small and large rectangles in the MULTI condition. If attention could not be divided and the large and small rectangles would be encompassed in a large unitary focus, enhanced activity in the medium ROI would be expected in the MULTI compared to the SINGLE conditions. Conversely, if attention could be exclusively allocated to the small and large rectangles without deploying attention to intervening regions, medium ROI responses would be similar across all conditions. The lack of differences between conditions in the first part of the RSVP (i.e., the first 20 s), indicated that in the MULTI condition no attention was allocated to the intervening regions in between the large and small rectangles. Although this suggested that attention was indeed initially divided in the MULTI condition, activity in the medium ROI was increased for the MULTI compared to SINGLE conditions in the second part of the RSVP. The sluggishness of the BOLD response did not appear to influence the difference between the early and late phase, since responses of the other ROIs did differ between conditions in the first part of the RSVP. Notably, the large and small ROI also showed an increased difference in the late compared to the early phase between the MULTI condition and the condition in which, respectively, the small and large rectangles were not attended. The increased activity in all three ROIs over the course of the MULTI condition seems to suggest that global arousal effects rather than changes in the attentional configuration might underlie the divergent medium ROI responses to the MULTI and SINGLE conditions across time. However, even if the highly complex attentional configuration can only be maintained for a limited time, the attentional distributions in the first part of the RSVP showed that it is possible to simultaneously deploy attention to non-contiguous regions in a highly complex configuration. The present findings are in agreement with results of an Event-Related Potential (ERP) study by Eimer (2000), which suggested that the annular region in between two simultaneously attended non-contiguous concentric ring-shaped regions of visual space could be ignored. However, the small size ($0.8^\circ \times 0.8^\circ$) and the similar features of the target (N) or distractor (M) letter that needed to be identified by the participants, might have promoted the use of a small attentional focus (or, alternatively, two foci) with high spatial resolution that did not cover the entire cued region but shifted from location to location within this region. The results reported by Eimer (2000), obtained by averaging over trials, did not provide insights concerning the

attentional distribution within trials. In contrast, in the present study the orientation of the rectangles could only be deduced by attending the *entire* stimulus. In addition, the color-orientation discrimination task could only be resolved by attending *both* stimuli, thereby reinforcing the use of an attentional configuration that encompassed both cued rectangles *on every trial*. Nevertheless, the results of Eimer (2000), together with findings of other psychophysical and neuroimaging studies (for review, see Chapter 6), showed that non-contiguous regions can be attended in parallel. The current study extends these findings by showing that these two regions can have complex shapes and can differ in size. Interestingly, recent findings suggest that dividing attention is not some exotic, rarely used capability of the attentional system, but rather might be an inherent feature of attentional allocation. By examining attentional modulations in monkey V1 when attention was shifted from one object to another, Khayat, Spekreijse, and Roelfsema (2006) showed that the increase in activity related to allocating attention to the new object preceded the suppression of activity corresponding to the disengagement of attention from the old object by ca. 60 ms. Thus, even during an apparently serial shift between objects, attention is engaged to the new object before it is disengaged from the old object, indicating that attention is allocated to both objects in parallel for a short period. Together, these results suggest that the distribution of attention is more appropriately characterized by a pliable landscape, in which the weights of multiple locations can be up- and down-regulated in parallel, than by a single unitary focus that moves from one location to another. Although the present findings in early visual areas cannot provide insights concerning the potential serial nature of later processing stages (e.g., object identification), they do indicate that attentional processing in early stages can occur in parallel.

Conclusion

The present preliminary results revealed that attention could be allocated in a more complex configuration than the single attentional focus suggested by the spotlight or zoom-lens metaphor. Attention could be allocated in an annular fashion and, moreover, attention could be divided over multiple non-contiguous regions of different sizes. No attention was deployed to intervening, irrelevant regions, indicating that the attentional system is highly efficient in allocating its limited resources even when highly complex configurations are required. This is in line with the emerging view that spatial attention mechanisms might be more appropriately characterized by parallel modulations in an “attentional landscape”, rather than by a serially moving unitary spotlight.

References

- Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance*, 26(2), 834-846.
- Bichot, N. P., Cave, K. R., & Pashler, H. (1999). Visual selection mediated by location: feature-based selection of non-contiguous locations. *Perception & Psychophysics*, 61(3), 403-423.
- Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nature Neuroscience*, 2(4), 370-374.
- Castiello, U., & Umiltà, C. (1992). Splitting focal attention. *Journal of Experimental Psychology: Human Perception and Performance*, 18(3), 837 - 848.
- Deichmann, R., Schwarzbauer, C., & Turner, R. (2004). Optimisation of the 3D MDEFT sequence for anatomical brain imaging: technical implications at 1.5 and 3 T. *Neuroimage*, 21, 757-767.
- Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, 369(6478), 313-315.
- Egley, R., & Homa, D. (1984). Sensitization of the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 10(6), 778 - 793.
- Eimer, M. (1999). Attending to quadrants and ring-shaped regions: ERP effects of visual attention in different spatial selection tasks. *Psychophysiology*, 36(4), 491-503.
- Eimer, M. (2000). An ERP study of sustained spatial attention to stimulus eccentricity. *Biological Psychology*, 52, 205-220.
- Eriksen, C. W., & James, D. S. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, 40(4), 225 - 240.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1-47.
- Hahn, S., & Kramer, A. (1998). Further Evidence for the Division of Attention Among Non-contiguous Locations. *Visual Cognition*, 5(1/2), 217-256.
- Juola, J. F., Bouwhuis, D. G., Cooper, E. E., & Warner, C. B. (1991). Control of attention around the fovea. *Journal of Experimental Psychology: Human perception and performance*, 17(1), 125-141.
- Khayat, P. S., Spekrijse, H., & Roelfsema, P. R. (2006). Attention lights up new object representations before the old ones fade away. *Journal of Neuroscience*, 26(1), 138-142.
- Kramer, A., & Hahn, S. (1995). Splitting the Beam: Distribution of Attention Over Non-contiguous Regions of the Visual Field. *Psychological Science*, 6(6), 381-386.
- McMains, S. A., & Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron*, 42(4), 677-686.
- McMains, S. A., & Somers, D. C. (2005). Processing efficiency of divided spatial attention mechanisms in human visual cortex. *Journal of Neuroscience*, 25(41), 9444-9448.
- Müller, M. M., & Hübner, R. (2002). Can the attentional spotlight be shaped like a doughnut? Evidence from steady state visual evoked potentials. *Psychological Science*, 13, 119 -124.
- Müller, M. M., Malinowski, P., Gruber, T., & Hillyard, S. A. (2003). Sustained division of the attentional spotlight. *Nature*, 424(6946), 309-312.
- Müller, M. M., Teder-Salejari, W., & Hillyard, S. A. (1998). The time course of cortical facilitation during cued shifts of spatial attention. *Nature Neuroscience*, 1(7), 631-634.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.
- Scharlau, I. (2004). Evidence for split foci of attention in a priming paradigm. *Perception & Psychophysics*, 66(6), 988-1002.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen, B. R., & Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268, 889-893.

- Shaw, M. L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal Experimental Psychology Human Perception and Performance*, 3, 201-211.
- Somers, D. C., Dale, A. M., Seifert, A. E., & Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academic Society USA*, 96(4), 1663-1668.
- Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. Stuttgart: Thieme.
- Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., & Dale, A. M. (1998). The retinotopy of visual spatial attention. *Neuron*, 21(6), 1409-1422.
- Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science*, 238(4828), 778-780.
- Wetherill, G. B. & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology*, 18, 1-10.
- Ward, R., Duncan, J., & Shapiro, K. (1996). The Slow Time-Course of Visual Attention. *Cognitive Psychology*, 30(1), 79-109.
- Woodman, G. F., & Luck, S. J. (2003). Serial deployment of attention during visual search. *Journal Experimental Psychology: Human Perception and Performance*, 29(1), 121-138.

Chapter 8

Summary

In this thesis, the networks of brain areas involved in controlling the goal-directed allocation of covert visual attention (*Chapters 2-5*) and the associated ‘top-down’ modulation of processing in the visual cortex (*Chapters 6-7*) were studied using functional magnetic resonance imaging (fMRI) and electro-encephalography (EEG). Moreover, to explore how proficiently goal-directed attention can be deployed in a range of situations, the studies presented in this thesis used various tasks which all required an efficient deployment of the limited attentional resources. These studies were part of three main projects, with each project focusing on a different aspect of voluntarily attentional control. The first project (*Chapter 2 & 3*) investigated whether task-irrelevant stimuli that match information in working memory (WM) can be ignored. The second project (*Chapter 4 & 5*) studied how accurately and efficiently the attentional focus can be scaled. Finally, the third project (*Chapter 6 & 7*) explored the limits of distributing attention in a complex spatial configuration. The results of these projects will be discussed within a broader context below.

A concise background on the current state of knowledge regarding the (neural) mechanisms of covert visual attention, as well as a short introduction to the neuroimaging methods used to study attention in this thesis, was presented in *Chapter 1*. In addition, the three projects that formed the basis of this thesis were introduced.

The first project (*Chapter 2 & 3*) incorporated psychophysical, Event-Related Potential (ERP), and fMRI measurements to explore the influence of WM content on visual search processes. WM plays an important role in visual search, since the representation of the searched object that is held in WM (the ‘search template’) guides attention to matching items in the visual scene. At a neurobiological level, this attentional guidance by WM content is implemented via top-down feedback from the prefrontal cortex that biases activity in visual areas in favor of matching visual inputs.

Because multiple items can be held in WM, an important question is whether all these items guide attention to matching stimuli, or only the items that are relevant for the task at hand. This question was examined using a paradigm in which participants had to search a target object in a stream of objects, while maintaining a second object in WM (the ‘memory item’) that would function as target in a subsequent task. The results showed that in the first task, the deployment of attention was exclusively guided by the search template and not by the ‘accessory’ memory item that was stored in WM for later use. In contrast to the search template, this accessory memory item did not bias attention to stimuli belonging to its object category. Moreover, when the memory item was presented as a distractor in the first task, this item (unlike the search target) did not receive more attention than other distractors that were not represented in WM. Conversely, the memory item *did* guide attention once it became relevant (being the new search target) in the second task, showing that the attentional guidance by items in

WM can be adjusted if a new target has to be searched. This dynamic control of the attentional guidance by WM content is presumably exerted by a subset of the frontal and parietal areas that are recruited when target information needs to be updated in between the first and second task. Although additional research is needed to assess how attentional guidance by WM representations is controlled within the identified network, the ERP and fMRI results of the first project indicated that the relevance of WM representations for the task at hand strongly influences whether they are allowed to provide top-down feedback. The finding that only relevant items in WM bias processing in visual areas (and that this relevancy can be adjusted according to task demands), is consistent with accumulating evidence indicating that the prefrontal cortex maintains information in a flexible and task-oriented manner rather than serving as a passive sensory buffer.

The second project tried to provide further insights into how the spatial attention focus (or ‘zoom-lens’) is scaled. Although attentional scaling is an important part of visual search, not much is known about its underlying mechanisms. To gain more insights into the characteristics of attentional scaling and its neural basis, we performed a series of four psychophysical studies (*Chapter 4*) and an additional rapid event-related fMRI study (*Chapter 5*). In a letter identification task, subjects were cued to attend to one of the three levels of a compound stimulus, consisting of three differently-sized superimposed letters. In ‘zoom’ trials, a letter-like symbol at the cued level indicated that attention had to be redirected to another level within the same stimulus. Response congruency between the different levels was varied to assess the degree of interference between levels when attention was scaled to a single level and, moreover, when attention needed to be rescaled to another level. Using this paradigm, primary functional properties of the scaling mechanism could be investigated by manipulating the rescaling ‘direction’ (i.e., zooming-in versus zooming-out) and by varying the required extent of rescaling.

The results showed that attention could be accurately scaled to each individual level. As expected, responses were slower and more error-prone when attention needed to be rescaled to another level. At the neural level, rescaling the zoom-lens recruited a fronto-parietal network, which overlapped with the network known to be involved in shifting attention from one location to another. The psychophysical studies suggested that both the extent as well as the direction of rescaling influenced the way in which the zoom-lens was adjusted. Performance decreased with an increase in the required extent of rescaling, and contracting appeared to be more difficult than expanding the zoom-lens. This latter difference between zooming-in and zooming-out was further explored in the fMRI study, which suggested that zooming-in and zooming-out were mediated by highly similar fronto-parietal networks. However, within these networks, areas related to cognitive or attentional control were activated more strongly during zooming-in. The decreased performance together with the increased need for attentional control

could suggest that contracting the zoom lens is a more difficult operation than expanding the zoom-lens. However, since differences in the visual context of the small and large letter in the used stimulus probably contributed to these divergent outcomes, additional studies with other stimuli are needed to further substantiate these findings.

Additionally, we addressed the question whether the intermediate level would be temporarily encompassed in the zoom-lens when attention was rescaled between the large and small level. Results of the psychophysical studies showed that conflicting information at the middle level deteriorated performance, when attention was rescaled between the large and small level (irrespective of the direction of rescaling). The negative effect of presenting response incongruent information at the middle level suggested that the zoom-lens could not be discretely rescaled from the large to the small level or the other way around. This interference was also reflected at the neural level by increased activation in the dorsal anterior cingulate cortex, which is involved in the detection of response conflict. However, these neural enhancements were only observed when attention was zoomed-out, suggesting that the middle level could be 'skipped' while zooming-in from the large to the small level. Although this latter finding is partially inconsistent with the results of the psychophysical experiments, one of these studies hinted to the possibility that the zoom-lens can be discretely rescaled with increased experience (although this effect was obtained for zooming-out rather than for zooming-in), which also could have influenced our fMRI results.

Overall, these divergent results indicate that attentional scaling mechanisms are influenced by both top-down factors, like strategy and practice, as well as bottom-up factors, such as visual context. Future studies could investigate the top-down and bottom-up factors influencing scaling mechanisms in more detail, by for example varying the amount of training and the type of stimulus.

The third project (*Chapter 6 & 7*) investigated whether spatial attention can be deployed in more complex configurations (e.g., attending multiple non-contiguous regions in parallel) than the single unitary focus suggested by the spotlight or zoom-lens model. It would be highly efficient if attention could be allocated to two discrete regions in parallel, thereby not wasting limited attentional resources on irrelevant intervening areas that would be encompassed by a large unitary focus. The issue whether attention can operate in a parallel fashion has been intensively debated in the literature. *Chapter 6* provides an overview of this debate and concludes, based on the combined evidence of psychophysical, neuroimaging, and modeling studies, that two discrete parts of a visual scene can be attended in parallel. However, the reviewed studies all used paradigms in which attention could be split in two foci with a similar size. Conversely, in the two psychophysical studies and the corresponding fMRI study presented in *Chapter 7*, attention needed to be simultaneously deployed to two complexly shaped regions that differed in size. Although preliminary, fMRI measurements of activity in the

visual cortex suggested that centrally aligned pairs of small and large outlined rectangles could be attended in parallel. No attention appeared to be deployed to the intervening regions, indicating that the attentional system is highly efficient in allocating its limited resources even when very complex configurations are required. In sum, the results of the third project support the emerging view that the distribution of attention is more appropriately characterized by a pliable landscape in which the weights of multiple locations can be modulated in parallel, than by a single serially moving focus.

Overall, the results of the discussed projects point to a rather proficient deployment of goal-directed attention in the explored situations. The first project showed that WM content exclusively guides attention to task-relevant objects in a visual scene and not to objects that will only become relevant in the near future. The findings of the second project indicated that the attentional focus can be adequately adjusted to match the spatial scale of a relevant region. However, the scaling of this focus appeared to occur in a continuous fashion by default, whereas more efficient discrete scaling only seemed to be possible after additional training. Finally, the third project revealed that spatial attention could be efficiently allocated in highly complex configurations. Although these results suggest that attention can be quite effectively deployed according to the behavioral goals of the observer, there are plenty of situations in which attention is more strongly controlled by stimulus-driven than by goal-directed factors. For example, top-down control often cannot prevent attentional capture by salient stimuli and, in addition, is sometimes not able to instantly override automatic responses (such as word reading in the Stroop task). Future research could examine the interaction between goal-directed and stimulus-driven attention by including salient stimuli in the used paradigms, leading to a more fine-grained delineation of the limitations of top-down controlled visual attention. Another interesting future line of research would be to study the control of goal-directed attention using a multi-method approach. Consistent with existing literature, the fMRI results of the first and second project suggest that brain areas in the frontal and parietal cortex control goal-directed attention in a variety of attention tasks. Combining fMRI with high temporal resolution methods like EEG might further elucidate whether frontal and parietal areas are conjointly recruited when attentional control is required or whether these areas are activated in succession. In addition, transcranial magnetic stimulation (TMS) can be applied to study causal relations between the areas of this fronto-parietal network. By inducing ‘virtual lesions’ with TMS, the functional roles of the recruited areas could be more clearly established, thereby providing a means to further deepen our understanding of the human visual attention system.

Samenvatting

Samenvatting

De studies in dit proefschrift onderzochten de netwerken in de hersenen die betrokken zijn bij het controleren van de doelgerichte sturing van coverte visuele aandacht (*Hoofdstuk 2-5*) en de daaraan gerelateerde *top-down* modulatie van verwerking in de visuele cortex (*Hoofdstuk 6-7*) m.b.v. functionele magnetische resonantie beeldvorming (fMRI) en electro-encefalografie (EEG). Om te onderzoeken hoe bekwaam doelgerichte aandacht gecontroleerd kan worden in uiteenlopende situaties, werden in de gepresenteerde studies verschillende soorten taken gebruikt die allemaal een efficiënte allocatie van de gelimiteerde aandachtsbronnen vereisten. De studies waren onderdeel van drie hoofdprojecten, die elk een eigen zwaartepunt hadden. In het eerste project (*Hoofdstuk 2 & 3*) werd onderzocht of irrelevante stimuli die overeenstemmen met representaties in het werkgeheugen genegeerd kunnen worden. De doelstelling van het tweede project (*Hoofdstuk 4 & 5*) was om erachter te komen hoe accuraat en efficiënt de omvang van de aandachtsfocus kan worden bijgesteld (oftewel, hoe accuraat en efficiënt aandacht kan worden “in- en uitgezoomd”). Tenslotte werd in het derde project bestudeerd hoe goed aandacht gedistribueerd kan worden in een complexe spatiële configuratie. In onderstaande paragrafen zullen de resultaten van deze projecten in een bredere context worden besproken.

Hoofdstuk 1 geeft een beknopt overzicht van de huidige kennis over de (neurale) mechanismen van coverte visuele aandacht. Daarnaast biedt het een korte inleiding in de beeldvormende methodes die in dit proefschrift gebruikt zijn om aandacht te bestuderen. In het laatste deel van hoofdstuk 1 worden de drie projecten geïntroduceerd, die beschreven worden in de overige hoofdstukken van het proefschrift.

In het eerste project (*Hoofdstuk 2 & 3*) werden psychofysische, gebeurtenis-gerelateerde potentialen (*Event-Related Potentials* of ERPs) en fMRI gegevens geanalyseerd om de invloed van werkgeheugen op visuele zoekprocessen in kaart te brengen. Werkgeheugen speelt een belangrijke rol tijdens zoeken, omdat het een representatie van het object dat gezocht wordt vasthoudt en deze ‘*search template*’ gebruikt om aandacht te richten op overeenkomende informatie in het visuele veld. Op een neurobiologisch niveau is deze sturing van aandacht door het werkgeheugen geïmplementeerd via een *top-down* feedback van de prefrontale cortex die activiteit in de visuele cortex beïnvloedt ten gunste van overeenkomende visuele input.

Omdat we meerdere items tegelijk in ons werkgeheugen kunnen vasthouden, is het een belangrijke vraag of al deze items aandacht sturen naar overeenkomende stimuli in het visuele veld, of alleen diegene die relevant zijn voor de uit te voeren taak. Deze vraag werd onderzocht m.b.v. een paradigma waarin proefpersonen een bepaald object moesten zoeken in een stroom van objecten, terwijl ze een tweede object moesten onthouden (het ‘*geheugen item*’) voor een volgende taak. De resultaten lieten zien dat in de eerste taak aandacht alleen werd gestuurd door de *search template* en niet door het geheugen item dat

in het werkgeheugen werd vastgehouden voor later gebruik. In tegenstelling tot de *search template* stuurde dit tweede item in het werkgeheugen aandacht niet in de richting van stimuli die tot zijn object categorie behoorden. Daarnaast kreeg dit geheugen item (in tegenstelling tot het gezochte object) niet meer aandacht dan andere objecten, als het in de eerste taak gepresenteerd werd. Daarentegen beïnvloedde het extra item in het geheugen de aandacht wél vanaf het moment dat het relevant werd (als het nieuwe object wat gezocht moest worden) in de tweede taak, wat laat zien dat het sturen van aandacht door werkgeheugenrepresentaties kan worden aangepast als er een nieuw object gezocht moet worden. Deze dynamische controle over welke representaties in het werkgeheugen aandacht mogen sturen, wordt waarschijnlijk uitgevoerd door een deel van de frontale en parietale gebieden die actief worden als informatie over het te vinden object wordt aangepast tijdens de overgang van de eerste naar de tweede taak. Hoewel verder onderzoek nodig is om inzicht te verkrijgen in de manier waarop deze hersengebieden de sturing van aandacht door het werkgeheugen precies controleren, laten de ERP en fMRI resultaten van het eerste project duidelijk zien dat de relevantie van een werkgeheugenrepresentatie voor de uit te voeren taak, bepaald of deze *top-down* feedback mag geven. De bevinding dat alleen relevante items in het werkgeheugen visuele verwerking beïnvloeden (en dat deze relevantie aangepast kan worden aan de hand van de taakstelling) is in overeenstemming met accumulerend bewijs dat aantoont dat de prefrontale cortex informatie vasthoudt op een flexibele, taakgerichte manier i.p.v. dat hij dient als passieve sensorische buffer.

Het tweede deelproject richtte zich op het bestuderen van de mechanismen die ten grondslag liggen aan het bijstellen van de omvang van de aandachtsfocus (“zoom-lens”). Hoewel het “inzoomen” en “uitzoomen” van aandacht belangrijk is tijdens het zoeken (op bijvoorbeeld een plattegrond), is nog niet veel bekend over de manier waarop dat gebeurt. Om meer inzicht te krijgen in dit “zoom” mechanisme en zijn neurale basis, hebben we vier psychofysische studies (*Hoofdstuk 4*) en een fMRI studie (*Hoofdstuk 5*) uitgevoerd. In een letter identificatie taak moesten proefpersonen letten op één van de drie niveaus van een stimulus die samengesteld was uit drie op elkaar liggende letters van verschillende grootte. In de zogenaamde ‘zoom’ trials werd op dat niveau in de stimulus een symbool gepresenteerd wat aangaf dat de proefpersoon binnen diezelfde stimulus, zijn of haar aandacht naar een ander niveau moest in- of uitzoomen. De respons congruentie tussen de verschillende niveaus werd gevarieerd om de mate van interferentie tussen de niveaus vast te stellen als aandacht werd gericht op een bepaald niveau, en wanneer aandacht vervolgens werd in- of uitgezoomd naar een ander niveau. Met dit paradigma konden we de belangrijkste eigenschappen van het “zoom” mechanisme bestuderen door het manipuleren van de “richting” (i.e., inzoomen versus uitzoomen) en de mate waarin de focus bijgesteld moest worden.

Samenvatting

De resultaten toonden aan dat aandacht accuraat op elk individueel niveau kon worden gefocust. Als aandacht vervolgens moest worden in- of uitgezoomd naar een ander niveau, werden proefpersonen zoals verwacht trager en maakte ze meer fouten. Het in- of uitzoomen naar een ander niveau activeerde in de hersenen een fronto-parietaal netwerk, dat overlap vertoonde met het netwerk wat betrokken is bij het verschuiven van de aandachtsfocus tussen twee locaties. De psychofysische resultaten suggereerden dat zowel de mate als de richting van het zoomen invloed hadden op hoe de focus werd bijgesteld. De prestatie verminderde als de focus meer moest worden bijgesteld en inzoomen was moeilijker dan uitzoomen. Deze laatste bevinding werd verder bestudeerd in een fMRI experiment, waarvan de resultaten suggereerden dat het in- en uitzoomen van aandacht fronto-parietale netwerken activeren die sterk op elkaar lijken. Binnen deze netwerken zijn gebieden die betrokken zijn bij controle processen echter sterker actief tijdens het inzoomen. De lagere prestatie van proefpersonen, gecombineerd met de sterkere behoefte aan cognitieve controle, zouden erop kunnen duiden dat inzoomen moeilijker is dan uitzoomen. Omdat verschillen in de visuele context van de grote en kleine letter in de gebruikte stimulus waarschijnlijk hebben bijgedragen aan deze resultaten, is echter vervolgonderzoek met andere stimuli nodig om deze resultaten verder te onderbouwen.

Daarnaast onderzochten we of de aandachtsfocus tijdelijk het middelste niveau omvat tijdens het in- en uitzoomen tussen het grote en kleine niveau. Resultaten van de psychofysische studies lieten zien dat als de focus wordt in- en uitgezoomd tussen het grote en kleine niveau, conflicterende informatie op het middelste niveau leidt tot een minder goede prestatie (ongeacht of in- of uitgezoomd wordt). Dit impliceert dat de zoom-lens niet discreet maar gradueel wordt bijgesteld. De interferentie van het middelste niveau werd op een neurobiologisch niveau gereflecteerd door verhoogde activatie in het dorsale gedeelte van de anterior cingulate cortex, een gebied dat betrokken is bij het detecteren van respons conflict. Deze verhoogde activering werd echter alleen waargenomen bij het uitzoomen, wat suggereert dat het middelste niveau wel kon worden “overgeslagen” tijdens het inzoomen van het grote naar het kleine niveau. Hoewel dit laatste resultaat gedeeltelijk in tegenstelling is met de bevindingen van de psychofysische experimenten, suggereerde een van deze studies dat de zoom-lens mogelijk discreet kan worden bijgesteld na extra training (al werd dit effect echter waargenomen bij het uit- in plaats van bij het inzoomen), iets wat ook onze fMRI resultaten beïnvloed zou kunnen hebben.

De uiteenlopende resultaten van de studies geven aan dat het bijstellen van de grootte van de aandachtsfocus beïnvloedt wordt door zowel *top-down* factoren, zoals strategie en training, als ook door *bottom-up* factoren, zoals visuele context. Toekomstige studies kunnen deze factoren verder bestuderen door bijvoorbeeld de hoeveelheid training en het type stimulus te variëren.

Het derde project (*Hoofdstuk 6 & 7*) onderzocht of spatiële aandacht toegepast kan worden in complexere configuraties (bijvoorbeeld de aandacht verdelen over meerdere gebieden die niet op elkaar aansluiten) dan een enkele focus zoals gesuggereerd door het *spotlight* of *zoom-lens* model. Het zou uitermate efficiënt zijn als aandacht verdeeld zou kunnen worden over twee discrete gebieden, zodat de beperkte aandachtscapaciteit niet verspild wordt aan irrelevante tussenliggende gebieden. Over de vraag of aandacht verdeeld kan worden is uitgebreid gedebatteerd in de literatuur. *Hoofdstuk 6* geeft een overzicht van dit debat en concludeert, gebaseerd op gecombineerd bewijs van psychofysische, *neuroimaging* en simulatie studies, dat aan twee discrete gebieden van het visuele veld tegelijk aandacht kan worden geschonken. De besproken studies gebruikten echter allemaal paradigma's waarin aandacht verdeeld kon worden over twee regio's met dezelfde grootte. In de twee psychofysische studies en de corresponderende fMRI studie die in *Hoofdstuk 7* gepresenteerd worden, moest aandacht echter verdeeld worden over twee gebieden met een complexe vorm, die ook nog in grootte verschilden. De voorlopige fMRI resultaten suggereren dat aandacht kon worden verdeeld over gecentreerde kleine en grote rechthoeken. Aan de tussenliggende regio's leek geen aandacht te worden gegeven, wat suggereert dat het aandachtssysteem buitengewoon efficiënt zijn gelimiteerde bronnen kan toepassen zelfs als dit om een zeer complexe aandachtsdistributie vraagt. De resultaten van het derde project ondersteunen het opkomende idee dat de distributie van aandacht beter beschreven kan worden als een glooiend landschap waarin de gewichten van meerdere locaties tegelijk kunnen worden gemoduleerd, dan als een enkele focus die serieel van locatie naar locatie verschuift.

De resultaten van de besproken projecten tonen aan dat doelgerichte aandacht adequaat kon worden toegepast in de onderzochte situaties. Het eerste project liet zien dat het werkgeheugen de aandacht exclusief richt op objecten in het visuele veld die nu belangrijk zijn en niet op objecten die pas later relevant worden. De bevindingen van het tweede project suggereerden dat de aandachtsfocus adequaat aangepast kan worden om overeen te stemmen met de omvang van een relevant gebied. De focus lijkt echter op een continue manier aangepast te worden, terwijl het efficiëntere discrete aanpassen alleen mogelijk leek na extra oefening. Tenslotte toonde het derde project aan dat spatiële aandacht efficiënt kon worden gedistribueerd in zeer complexe configuraties. Hoewel deze resultaten suggereren dat doelgerichte aandacht behoorlijk effectief toegepast en goed gecontroleerd kan worden, zijn er talrijke situaties waarin aandacht sterker gecontroleerd wordt door stimulus-gedreven (*stimulus-driven*) dan doelgerichte factoren. *Top-down* controle kan bijvoorbeeld niet altijd voorkomen dat aandacht getrokken wordt door opvallende stimuli en is ook niet altijd in staat om automatische reacties te onderdrukken (zoals het lezen van woorden tijdens de *Stroop* taak). Toekomstig onderzoek zou de interactie tussen doelgerichte en stimulus-gedreven aandacht kunnen onderzoeken door aandachttrekkende stimuli in de gebruikte paradigma's

Samenvatting

te implementeren, waardoor we de limitaties van doelgerichte visuele aandacht verder in kaart zouden kunnen brengen. Het bestuderen van doelgerichte aandacht met een combinatie van methoden zou een andere interessante onderzoekslijn vormen. De fMRI resultaten van het eerste en tweede project suggereerden dat, in overstemming met de bestaande literatuur, hersengebieden in de frontale en parietale cortex doelgerichte aandacht controleren in verschillende aandachtstaken. Het combineren van fMRI met methoden die een hoge temporele resolutie hebben (zoals EEG) zou duidelijk kunnen maken of de verschillende gebieden van dit fronto-parietale netwerk tegelijkertijd of na elkaar geactiveerd worden als aandacht moet worden verschoven. Daarnaast kunnen de causale relaties tussen deze frontale en parietale gebieden bestudeerd worden m.b.v. transcraniële magnetische stimulatie (TMS). Door het induceren van ‘virtuele laesies’ met TMS, zouden de functies van de betrokken gebieden met meer zekerheid vastgesteld kunnen worden, wat ons de mogelijkheid biedt om ons inzicht in de werking van het menselijke visuele aandachtssysteem verder te vergroten.

Acknowledgments

Eindelijk is het zo ver, mijn proefschrift is af! Op dit belangrijke hoofdstuk na dan, waarin ik alle mensen wil bedanken die hebben geholpen dit proefschrift tot stand te brengen.

Rainer and Pieter, I am honored and grateful to be supervised by two excellent researchers with such an inspiring and stimulating passion for research. I deeply appreciate your support and your trust in my scientific abilities. Rainer, your contagious enthusiasm inspired me and kept me motivated, whereas your pragmatism kept me focused – when necessary you adequately scaled my attentional focus between zooming-out on all the intriguing research questions waiting to be studied and zooming-in on the projects that needed to be finished. Pieter, ik bewonder je visie en je gedreven manier van onderzoek doen. Ik heb je snelle, zeer nuttige feedback op mijn manuscripten heel erg gewaardeerd: als een uitermate effectieve search template richtte je mijn aandacht op de relevante punten van het eerste project, precies dat wat ik nodig had als andere *accessory* items me – geheel tegen ons model in – afleidde. Peter, mijn ‘derde begeleider’, wat ben ik blij dat je naar Maastricht bent gekomen! Samen onderwijs verzorgen was niet alleen leerzaam, maar vooral ook erg leuk. Daarnaast zijn je bijdragen aan mijn projecten (vooral de laatste) van onschatbare waarde.

Bettina en Ellen, ik ben erg blij dat jullie als paranimfen naast me willen staan tijdens de promotie. Bettina, bedankt voor de hulp met de fMRI analyses (je bent en blijft mijn hersen解剖学-TomTom), de support, lachbuien, discussies, *Kartoffelpuffer* en vriendschap - je bent de ideale kamergenoot. Ellen, de lunches, de yells, en de peptalks, ik heb ze erg gewaardeerd – echt leuk dat we nu in dezelfde groep werken.

Many other (ex-)colleagues of Maastricht University made going to work a joy: Alard, Alex, Aline, Amanda, Annemie (bedankt voor het geregeld), Anniek, Bernadette (thanks for the support), Bert (bedankt voor de leuke discussies), Chantal (ik kijk uit naar onze samenwerking), Elia, Federico, Francesco (2x), Fren, Giancarlo, Hanneke, Henk, Holger, Ingrid, Lars, Marin, Mart, Martien (merci voor de steun), Michael, Milene, Niels, Nienke, Niko, Nina, Noël, Petra, Riny, Sandra, Sven, Teresa, Valerie, Vera, Vincent and all others. Additional thanks to the BrainInnovation crew for methodological support, and to Stefan Debener & Alex Strobel for showing me the ins & outs of combined EEG-fMRI measurements. De NWO projectgroep, bedankt voor de goede discussies tijdens onze meetings. Jan en Martijn, bedankt voor het coördineren van het NWO project en Jan, bedankt voor het lezen van mijn proefschrift. Gwendid, onze funny e-conversations, briljante onderzoeks- en bedrijfsplannen, en gezellige bijkletsdates waren het perfecte medicijn tegen onderzoeksfrustraties: ik ben blij dat ik je heb leren kennen tijdens mijn promotietijd.

Acknowledgments

Lieve vriendinnen, Pleun en de B.E.B.ers (Corlijn, Edith, Elske, en Tamara), bedankt dat jullie me af en toe het lab uit - en de kroeg in - hebben weten te praten. Dat jullie allemaal weer in de randstad (en dan ook nog in Den Haag!) zijn gaan wonen was natuurlijk niet de bedoeling, maar gelukkig weten we elkaar nog altijd te vinden!

Lieve Marja, bedankt voor al je liefde, steun en interesse door de jaren heen, je bent een supermoeder! Dit proefschrift is opgedragen aan jou, jij bent mijn basis. Onze (nachtelijke) discussies over de mysteries van het brein waren de oorsprong van mijn interesse in *neuroscience*. Dat we een andere invalshoek hebben maakt de discussie alleen nog maar interessanter. Lieve Hans, Wally, Dorette, Marcel, Mirko (heel erg bedankt voor het ontwerpen van de prachtige kaft!), Rebecca en Dorotheé (bedankt voor de literatuur updates) wat ben ik blij dat ik er zo'n leuke tweede familie bij heb gekregen!

Lieve, lieve, lieve, Joel, dit proefschrift is eigenlijk ook jouw proefschrift. Je hebt me echt bij elke stap in het onderzoek geholpen: van meedenken met het design, geduldig proefpersoon zijn bij de vele testmetingen en suggesties voor de data-analyses, tot feedback op mijn manuscripten en hulp bij het maken van de figuren (het belang van een hoog IQ [IllustratorQuotiënt] kan nooit overschat worden!) – ondertussen mijn neurotisch gestress (dat bij elke stap toenam..) sussend. En dat terwijl je tegelijkertijd ook met het afronden van je eigen proefschrift bezig was! Meer nog dan je praktische hulp en je zowel motiverende als relativerende peptalks, ben ik je eeuwig dankbaar voor al je steun, liefde.. voor alles eigenlijk – iets waarvoor bedanken simpelweg niet volstaat. Samen aan onze proefschriften werken in onze eigen *dataprocessingroom* annex woonkamer was heel gezellig, maar misschien wordt het nu tijd voor een computerloze vakantie?

Curriculum Vitae

Judith Peters was born on September 9th 1979 in Leiden, The Netherlands. In 1997 she completed her secondary education at the Dalton Scholengemeenschap in The Hague, after which she studied Psychology at Maastricht University. During her study, Judith worked as a teaching assistant and research assistant. After an internship under supervision of prof. dr. R. Goebel and dr. E. Formisano, she received her master's degree in biological psychology (neuropsychology) in 2002. In the same year Judith started her PhD-project at the Cognitive Neuroscience Department of Maastricht University, resulting in this thesis. Currently, she continues her work at this department as a postdoctoral fellow.